

ΑΤΕΙ ΜΕΣΟΛΟΓΓΙΟΥ  
ΤΜΗΜΑ ΕΦΑΡΜΟΓΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗ ΔΙΟΙΚΗΣΗ ΚΑΙ ΣΤΗΝ  
ΟΙΚΟΝΟΜΙΑ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΤΙΤΛΟΣ ΕΡΓΑΣΙΑΣ:

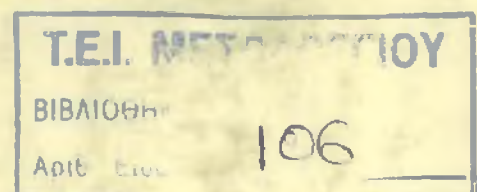
ΠΡΩΤΟΤΥΠΟΠΟΙΗΣΕΙΣ ΓΙΑ ΚΩΔΙΚΟΠΟΙΗΣΗ ΚΑΙ  
ΜΕΤΑΔΟΣΗ ΠΟΛΥΜΕΣΙΚΩΝ ΔΕΔΟΜΕΝΩΝ -  
ΠΡΩΤΟΚΟΛΛΑ ΠΡΑΓΜΑΤΙΚΟΥ ΧΡΟΝΟΥ

ΕΠΙΒΛΕΤΩΝ ΚΑΘΗΓΗΤΗΣ:  
Δρ. ΝΙΚΟΛΑΟΣ ΚΟΣΜΑΣ

ΦΟΙΤΗΤΕΣ:  
ΑΘΑΝΑΣΟΥΛΗΣ ΔΗΜΗΤΡΙΟΣ Α.Μ.9596  
ΛΑΜΠΙΡΗΣ ΒΑΣΙΛΕΙΟΣ Α.Μ.8917

ΜΕΣΟΛΟΓΓΙ 2005

<b>ΜΕΡΟΣ Α:</b> Πρωτοτυποποιήσεις για κωδικοποίηση και μετάδοση multimedia δεδομένων.....	6
1. Εισαγωγή.....	6
1.1 Τι είναι τα multimedia.....	6



ΑΤΕΙ ΜΕΣΟΛΟΓΓΙΟΥ  
ΤΜΗΜΑ ΕΦΑΡΜΟΓΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗ ΔΙΟΙΚΗΣΗ ΚΑΙ ΣΤΗΝ  
ΟΙΚΟΝΟΜΙΑ

*ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ*

ΤΙΤΛΟΣ ΕΡΓΑΣΙΑΣ:

**ΠΡΩΤΟΤΥΠΟΠΟΙΗΣΕΙΣ ΓΙΑ ΚΩΔΙΚΟΠΟΙΗΣΗ ΚΑΙ  
ΜΕΤΑΔΟΣΗ ΠΟΛΥΜΕΣΙΚΩΝ ΔΕΔΟΜΕΝΩΝ -  
ΠΡΩΤΟΚΟΛΛΑ ΠΡΑΓΜΑΤΙΚΟΥ ΧΡΟΝΟΥ**

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:  
Δρ. ΝΙΚΟΛΑΟΣ ΚΟΣΜΑΣ

ΦΟΙΤΗΤΕΣ:  
ΑΘΑΝΑΣΟΥΛΗΣ ΔΗΜΗΤΡΙΟΣ Α.Μ.9596  
ΛΑΜΠΙΡΗΣ ΒΑΣΙΛΕΙΟΣ Α.Μ.8917

ΜΕΣΟΛΟΓΓΙ 2005

<b>ΜΕΡΟΣ Α:</b> Πρωτοτυποποιήσεις για κωδικοποίηση και μετάδοση multimedia δεδομένων.....	6
1.Εισαγωγή.....	6
1.1 Τι είναι τα multimedia.....	6

1.2 Η ανάγκη για χειρισμό πολύπλοκων τύπων δεδομένων .....	7
2. Εικόνα .....	9
2.1. Εικόνες και Εφαρμογές .....	9
2.2 Πρότυπο JPEG .....	9
2.2.1. Βασικά Πλεονεκτήματα Χρήσης JPEG .....	10
2.3 JPEG Χωρίς Απώλειες .....	11
2.4 Πρότυπο JBIG .....	13
2.5 Πρότυπο GIF.....	13
2.5.1 Προβλήματα GIF .....	14
2.5.2 Οργάνωση του Αρχείου .....	14
2.5.3 GIF87a .....	15
2.6 Fractal Image Compression (Απειροστική Συμπύεση Εικόνας).....	15
2.6.1 Γενικά στοιχεία σχετικά με την απειροστική συμπύεση εικόνας .....	15
2.6.2 Λειτουργία απειροστικής συμπύεση εικόνας.....	16
2.6.3 Απώλεια κλίμακας και αύξηση της ανάλυσης.....	16
2.7 Απειροστικότητα – JPEG .....	18
2.8 JPEG –GIF .....	18
3. Ήχος .....	20
3.1 Ήχος και Εφαρμογές .....	20
3.2 Συμπύεση .....	20
Μουσική και υπολογιστές .....	20
3.3 Πρότυπα ITU .....	22
3.3.1 Οπτικοακουστική Τηλεφωνία (Audiovisual Telephony).....	22
• G.721.....	22
• G.722.....	22
• G.723.....	23
• G.728.....	23
3.4 Τα πρότυπα MPEG .....	23
• MPEG-1 .....	24
• MPEG-2 .....	25
4. Video .....	26
4.1 Video και Εφαρμογές .....	26
4.2 Σύλληψη Video .....	26
4.3 Τεχνικές Συμπύεσης .....	27
4.4 Το πρότυπο συμπύεσης MPEG .....	27
• MPEG-1 Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s .....	28
• MPEG-2 Generic coding of moving pictures and associated Audio information .....	28
• MPEG-3 .....	28
• MPEG-4 Coding of audio-visual objects .....	29
• MPEG-7 .....	29
4.5 Είδη πλαισίων (frames) .....	30
• Ενδο-πλαισιακή Κωδικοποίηση (Intra-frame Coding).....	30

• Δια-πλαισιακή Κωδικοποίηση (Inter-frame Coding)	30
Σε ένα σήμα (bit-stream) MPEG υπάρχουν τριών ειδών πλαίσια	30
• I (Intra frames)	30
• P (Predicted frames)	31
• B (Bi-directional frames)	31
4.6 Ψηφιακό interactive βίντεο (Digital video interactive DVI)	33
4.7 Πρότυπα ITU-T H.200	34
• H.200	35
• H.221	35
• H.222	36
• H.230	36
• H.242	36
• H.231 και H243	37
• H.261	37
• H.320	38
5. Μετάδοση Multimedia δεδομένων	39
5.1 Διαδικασία δικτύου	39
5.2 Μετάδοση πολυμέσων	39
5.3 Πρότυπα για μετάδοση βίντεο	39
• H.261	39
• H.233	40
• H.320	40
• H.263	41
ΜΕΡΟΣ Β: Πρωτόκολλα πραγματικού χρόνου	42
1. ΕΙΣΑΓΩΓΗ	42
1.1 ΕΞΕΛΙΞΗ ΤΟΥ ΔΙΚΤΥΟΥ	42
1.1.1 Δεκαετία του 60. Η εκκίνηση	42
1.1.2 Δεκαετία του 70. οι πρώτες συνδέσεις	42
1.1.3 Δεκαετία του 80. ένα παγκόσμιο δίκτυο για την ακαδημαϊκή κοινότητα	43
1.1.4 Δεκαετία 90. ένα παγκόσμιο δίκτυο για όλους	43
2. Multimedia Over IP:	44
2.1 Multimedia networking: Στόχοι και προοπτικές	44
2.2 Επικοινωνία πραγματικού χρόνου	44
2.3 Multimedia και Internet	45
3. Γιατί τα παραδοσιακά πρωτόκολλα μεταφοράς δεν είναι κατάλληλα για multimedia networking	46
3.1 Δομή των πρωτοκόλλων	46
3.2 Πολύπλεξη	46
3.3 Έλεγχος ροής	47
3.4 Έλεγχος λαθών	48
3.5 Πληροφορίες ελέγχου	48
3.6 Έλλειψη QOS χαρακτηριστικών	49

4. Χαρακτηριστικά που πρέπει να έχουν τα πρωτόκολλα πραγματικού χρόνου .....	50
4.1 Υψηλό throughput .....	50
4.2 Δυνατότητα multicast .....	51
4.3 Χαρακτηριστικά QOS .....	51
5. Πρωτόκολλα δέσμευσης πόρων .....	52
5.1 RSVP (Resource ReSerVation Protocol) .....	52
5.1.1 Ανάπτυξη .....	53
5.1.2 Χαρακτηριστικά του RSVP .....	53
5.1.3 Πως δουλεύει το RSVP .....	54
5.2 ST-II .....	56
5.2.1 Μορφή ST-II σύνδεσης .....	56
5.2.2 Χαρακτηριστικά του stream που μεταδίδει το ST-II .....	58
5.3 Σύγκριση RSVP και ST-II .....	59
6. Πρωτόκολλα Μεταφοράς .....	61
6.1 RTP/RTCP (Real-time Transport Protocol/Real-Time Control Protocol).61	
6.1.1 Πως λειτουργεί το RTP .....	62
6.1.2 RTP fixed header fields .....	63
6.2 RTCP (Real-Time Control Protocol) .....	64
6.3 Χαρακτηριστικά του RTP/RTCP .....	64
7. Πρωτόκολλο streaming .....	66
7.1 Το πρωτόκολλο RTSP (Real Time Streaming Protocol) .....	66
7.1.1 Γενική περιγραφή.....	66
7.1.2 Αναλυτικά οι αιτίες ύπαρξης του πρωτοκόλλου RTSP .....	66
7.1.3 Λειτουργίες του πρωτοκόλλου .....	67
7.1.4 Ιδιότητες του πρωτοκόλλου .....	68
7.1.5 Η γενική λειτουργία του πρωτοκόλλου .....	70
7.1.6 Καταστάσεις του RTSP .....	71
7.1.7 Σχέση με άλλα πρωτόκολλα .....	71
7.1.8 Συνδέσεις .....	72
7.1.8.1 Pipelining .....	73
7.1.8.2 Αξιοπιστία και βεβαιώσεις λήψης .....	73
8. Πρωτόκολλο SIP (Session Initiation Protocol) .....	75
8.1 Ιστορία SIP .....	75
8.2 Λειτουργία SIP .....	77
8.3 Τρόπος λειτουργίας SIP .....	77
8.4 Πλεονεκτήματα SIP .....	79
Υπηρεσίες .....	79
Εξελιξιμότητα .....	79
Επεκτασιμότητα.....	80
Ευελιξία .....	80
8.5 Παροχή υπηρεσιών SIP .....	80
8.6 Ποιότητα Υπηρεσίας SIP .....	81
8.7 Μελλοντικές κατευθύνσεις .....	82

9.Βιβλιογραφία .....	84
<b>Σχήμα 1.</b> Ηχητική ποιότητα και μέθοδος ψηφιοποίησης .....	21
<b>Σχήμα 2.</b> Πρότυπα για κωδικοποίηση ήχου .....	22
<b>Σχήμα 3.</b> Αναλογία πλαισίων I,P,B, σε ένα σήμα MPEG .....	32
<b>Σχήμα 4.</b> Οπτικοακουστικές υπηρεσίες της ITU τυποποίησης .....	36
<b>Σχήμα 5.</b> Δέσμευση σε ένα κόμβο του μονοπατιού μετάδοσης των δεδομένων .....	54
<b>Σχήμα 6.</b> Η ενοποίηση των αιτήσεων δέσμευσης .....	55
<b>Σχήμα 7.</b> Δέσμευση πόρων με το ST-II .....	58
<b>Σχήμα 8.</b> Επικεφαλίδα της ST-II PDU.....	59
<b>Σχήμα 9.</b> RTP δεδομένα σε ένα IP πακέτο .....	62
<b>Σχήμα10.</b> RTP Packet Header .....	63

# ΜΕΡΟΣ Α: Πρωτοτυποποιήσεις για κωδικοποίηση και μετάδοση multimedia δεδομένων

## 1. Εισαγωγή

---

### 1.1. Τι είναι τα multimedia

Τα πολυμέσα (multimedia) είναι πιθανόν ένας από τους πιο χρησιμοποιημένους όρους της δεκαετίας του ενενήντα[3]. Το πεδίο είναι μία διασταύρωση από πέντε σημαντικές βιομηχανίες: υπολογισμού τηλεπικοινωνιών, εκδοτικής, ηλεκτρονικών βίντεο και ήχου και εκπομπής τηλεόρασης και κινηματογραφικών ταινιών. Η εμφάνιση στην δεκαετία του εβδομήντα ενός τεχνολογικού πεδίου, της δικτύωσης υπολογιστών, οδήγησε στην πρώτη φάση της ενεργής συνεργασίας μεταξύ του υπολογισμού και των τηλεπικοινωνιών.

Τα πολυμέσα δεν έφεραν μόνο τρεις νέες βιομηχανίες στο χώρο αλλά πρόσθεσαν και μία νέα διάσταση στη δυναμική αγορά: όσο η δικτύωση υπολογιστών στοχεύει ουσιαστικά σε μία επαγγελματική αγορά, τα πολυμέσα αφορούν και εμπορικούς και καταναλωτικούς τομείς. Έτσι, η σύνθεση της αγοράς τηλεπικοινωνιών δεν είναι μόνο επαγγελματικά και βιομηχανικά δίκτυα ή υψηλής ταχύτητας μισθωμένα κυκλώματα ή συνεταιρικά δίκτυα δεδομένων, αλλά καλύπτει επίσης και την πρότυπη τηλεφωνία ή χαμηλής ταχύτητας Ολοκληρωμένες Υπηρεσίες Ψηφιακών (Integrated Services Digital Networks, N-ISDN). Ομοίως, δεν εμπλέκεται μόνο η επαγγελματική αγορά βίντεο και ήχου, αλλά και η αντίστοιχη καταναλωτική όπως και η σχετική τηλεόραση, οι κινηματογραφικές ταινίες, και ο τομέας εκπομπής.

Έτσι τελικά δεν είναι έκπληξη ότι τα πολυμέσα βρίσκουν δυσκολίες στο να αποφύγουν ασάφειες στο σκοπό την πολλαπλότητα των ορισμών και στην ασταθή ορολογία, για να καθιερωθούν.

Όταν στις αρχές της δεκαετίας του ενενήντα οι επαγγελματίες

Αναρωτιόνταν τι είναι τα πολυμέσα, οι περισσότεροι έπρεπε να παραδεχτούν τη δυσκολία να δώσουν κάποιο ορισμό. Επαγγελματικά περιοδικά συχνά αναγνωρίζουν την ασάφειά τους: «Τα πολυμέσα είναι εξ ορισμού απροσδιόριστα» ή «αν ρωτήσεις 10 διαφορετικά άτομα να σου δώσουν ένα ορισμό για τα πολυμέσα, τότε σίγουρα θα πάρεις 10 διαφορετικές απαντήσεις» αναφέρει ο G.R. Wichman το Δεκέμβριο του 1991 (Wichman, 1991). Με την εφαρμογή σε εξειδικευμένα περιβάλλοντα, η δυσκολία στον ορισμό τους παραμένει: «Στο χώρο εργασίας η έννοια είναι θολή», αναφέρει ο A. Speed (Speed, 1991).

## 1.2 Η ανάγκη για χειρισμό πολύπλοκων τύπων δεδομένων

Τα αντικείμενα πολυμέσων επιβάλλουν αρκετές λειτουργικές απαιτήσεις στην υποδομή της επεξεργασίας δεδομένων σαν αποτέλεσμα της ανάγκης για χειρισμό πολύ μεγάλων και πολύπλοκων τύπων δεδομένων. Αυτά παρουσιάζουν ειδικά θέματα αναπαράστασης, παραποίησης, διαχείρισης και αποθήκευσης δεδομένων. Το παλαιάς τεχνολογίας υλικό και λογισμικό και τα συστήματα πελάτη-εξυπηρετητή (server-client) σήμερα δεν είναι σχεδιασμένα για να χειρίζονται μαζικούς τύπους multimedia [4]δεδομένων και πρέπει να εμπλουτιστούν με επιπρόσθετο υλικό και λογισμικό ή να αντικατασταθούν με νέες συσκευές με ενσωματωμένες multimedia δυνατότητες.

Η μετάδοση και αποθήκευση τέτοιων συμπαγών και μη δομημένων δεδομένων, ειδικά σε δικτυωμένες εφαρμογές πολλών χρηστών, μπορούν να δημιουργήσουν ταινίες ασφαλείας σε πολλά υπάρχοντα συστήματα των οποίων η χωρητικότητα του εύρους ζώνης είναι συχνά ανεπαρκής να χειρίζεται εντάσεις οι οποίες συχνά λειτουργούν σε mode πραγματικού χρόνου.

Σαν αποτέλεσμα, υπάρχει ανάγκη για hardware και software λύσεις οι οποίες πρέπει να περιέχουν χαρακτηριστικά για χειρισμό μετάδοσης και αποθήκευσης συμπαγών multimedia δεδομένων χωρίς να βλάπτουν την απόδοση του όλου συστήματος.

Οι interactive εφαρμογές πολυμέσων δεν μπορούν χωρίς τη συμπίεση και την αποσυμπίεση στην ανταλλαγή πολυμέσων. Το υπαρκτό υλικό και τα δίκτυα μπορεί να έχουν επαρκές εύρος Ζώνης για να χειριστούν τη μετάδοση βίντεο, αλλά αυτές οι δυνατότητες μπορούν να μελετούνται μαζί με όλες τις άλλες διακινήσεις στα δίκτυα LAN και WAN το οποία πρέπει να εξυπηρετούνται ταυτόχρονα. Αυτά τα δίκτυα ραγδαία έρχονται στον κορεσμό ακόμη και χωρίς τη διακίνηση πολυμέσων καθώς ο αριθμός των χρηστών και η διακίνηση δεδομένων αυξάνεται σε σημείο που η απαίτηση χρόνου γίνεται αφόρητη.

Τα πολυάριθμα σχέδια συμπίεσης και αποσυμπίεσης υπάρχουν και αυτό κάνει την επεξεργασία των πολυμέσων μια πρακτική πραγματικότητα χρησιμοποιώντας υλικό λογισμικό ή και συνδυασμό των δύο. Τα δεδομένα βίντεο μπορούν να συμπεστούν σε πραγματικό χρόνο αφού πρώτα συλληφθούν και ψηφιοποιηθούν ή αφού αποθηκευθούν σε ένα σύστημα. Η αποθήκευση συνεπάγεται ότι επαρκής και γρήγορη αποθήκευση ικανότητα είναι διαθέσιμη μέσα στο σύστημα.

Η συμπίεση βίντεο εκμεταλλεύεται το γεγονός ότι το ανθρώπινο μάτι είναι λιγότερο ευαίσθητο στις διακυμάνσεις του χρώματος και των σχημάτων από ότι στη λάμψη της εικόνας.



Η συμπίεση συνεπώς βασίζεται στην επινόηση αλγορίθμων οι οποίοι συνοψίζουν μέσα στο ψηφιακό βίντεο όλες τις λεπτομέρειες οι οποίες δε γίνονται αντιληπτές κανονικά από το ανθρώπινο μάτι. Αυτό επιτρέπει σε κάθε πλαίσιο βίντεο να περιέχει λιγότερα δεδομένα δηλαδή μειώνεται το μέγεθος της αποθήκευσης, επιταχύνεται η μετάδοση της εικόνας, και οδηγεί σε συμπίεση των πλαισίων.

Υπάρχουν δύο φόρμες συμπίεσης που χρησιμοποιούνται σήμερα. Η συμμετρική συμπίεση παίρνει τον ίδιο χρόνο για συμπίεση και αποσυμπίεση. Η ασυμμετρική συμπίεση παίρνει περισσότερο χρόνο για να κωδικοποιήσει αλλά συνήθως οδηγεί σε υψηλά ποσοστά συμπίεσης και σε καλύτερη ποιότητα εξόδου. Υπάρχουν επίσης ειδικά σχέδια συμπίεσης χρησιμοποιώντας κλασματικά αντικείμενα το οποία θεωρητικά μπορούν να φτάσουν ποσοστά συμπίεσης της τάξης 10000:1 αλλά παίρνουν πολύ χρόνο για να διεκπεραιωθούν ώστε να αποκτήσουν αξία σε απρόβλεπτες μεταδόσεις όπως οι διασκέψεις πολυμέσων πραγματικού χρόνου.

Η βασική λειτουργία της ψηφιακής συμπίεσης είναι η αναγνώριση και η εξάλειψη των περιττών δεδομένων μέσα αλλά και μεταξύ των διαδοχικών πλαισίων, αυτό ως αποτέλεσμα τη μείωση των δεδομένων που πρέπει να μεταδοθούν. Το ανθρώπινο μάτι έχει δυσκολία στο να αντιληφθεί τις αλλαγές από το ένα πλαίσιο στο άλλο. Η εξάλειψη των περιττών δεδομένων μεταξύ των πλαισίων και η κωδικοποίηση μόνο των

αλλαγών είναι γνωστή ως interframe συμπίεση, η οποία επιτρέπει πολύ υψηλότερα ποσοστά συμπίεσης. Η interframe συμπίεση συγκρίνει τα δεδομένα στα συνεχόμενα πλαίσια ενώ η interframe συμπίεση αναλύει τις παρόμοιες εμφανιζόμενες περιοχές μέσα σε ένα πλαίσιο. Και οι δύο μέθοδοι χρησιμοποιούνται στα σχέδια συμπίεσης όπως τα JPEG, MPEG και το ψηφιακό interactive βίντεο (DVI).

Τα σχέδια συμπίεσης διακρίνονται σε lossy και lossless και αυτό εξαρτάται από το αν όλα τα αρχικά δεδομένα αποκαθίστανται μετά την αποσυμπίεση. Οι περισσότερες συμπίεσης είναι της μορφής lossy αλλά το μέγεθος των χαμένων δεδομένων ποικίλει ανάλογα με το ρυθμό των πλαισίων, το τελικό μέγεθος της εικόνας, το βάθος χρώματος, την εμβέλεια των συχνότητα ήχου αλλά και από παράγοντες όπως η αδρότητα και η αντίθεση της εικόνας. Αυτό σημαίνει ότι πολλές ανταλλαγές μπορούν να γίνουν μεταξύ της αναλογίας συμπίεσης και της αποδεκτής ποιότητας. Ο κώδικας συμπίεσης και αποσυμπίεσης μπορεί να υλοποιηθεί χρησιμοποιώντας ειδικά DSP μικροτσίπ, ρουτίνες λογισμικού, ή και συνδυασμό και των δύο. Ο χρόνος επεξεργασίας του κώδικα είναι καθοριστικό σημείο στην υπολογισμό της καθυστέρησης ή της λανθάνουσας κατάστασης στις μεταδόσεις interactive βίντεο.

Η ύπαρξη τους συνεισφέρει στην καθυστέρηση κατά τη διάρκεια της μετάδοσης, της λήψης και σε άλλα στάδια όπου η διαδικασία συμπίεσης πρέπει να λάβει χώρα. Αυτή η καθυστέρηση δεν είναι πολύ καθοριστική στις εφαρμογές βίντεο αποθήκευσης και προώθησης, αλλά μελέτες έχουν δείξει ότι αθροιστικά η καθυστέρηση δεν θα πρέπει να ξεπερνά τα 100 ms σε ένα μονό κανάλι για ιδανική λειτουργία.

## 2 Εικόνα

### 2.1 Εικόνες και Εφαρμογές

Η εικόνα έχει γίνει απαραίτητο στοιχείο κάθε σύγχρονης. Ακόμα και σε περιπτώσεις όπου η εικόνα δεν αποτελεί αντικείμενο της εφαρμογής, οι απαιτήσεις για απλά και κατανοητά interfaces εισάγουν αναπόφευκτα την εικόνα. Υπάρχουν διάφορα είδη εικόνας, κάθε ένα εκ των οποίων είναι κατάλληλο για ορισμένα είδη εφαρμογών.

Το πιο απλό, σε σχέση με την πολυπλοκότητα της απεικόνισης του στον υπολογιστή, είναι οι διτονικές (bitonal) εικόνες[5]. Χαρακτηριστικό αυτής της κατηγορίας είναι η ύπαρξη μόνο δύο χρωμάτων (μαύρο και άσπρο συνήθως). Οι διτονικές εικόνες βρίσκουν εφαρμογή σε προγράμματα οργάνωσης επιχειρήσεων και οργανισμών όπου παρουσιάζεται η ανάγκη αρχειοθέτησης εγγράφων, αποδείξεων, επιταγών κ.λπ. Αυτές οι εικόνες προέρχονται από scanning των εγγραφών και αποθηκεύονται σε ειδικού σκοπού συστήματα διαχείρισης βάσεων δεδομένων. Άλλες περιπτώσεις όπου παρουσιάζονται συχνά διτονικές εικόνες είναι τα τεχνικά σχέδια, τα διαγράμματα, οι χάρτες κ.λπ.

Στο επόμενο επίπεδο έχουμε τις εικόνες συνεχούς τόνου (continuous tone images). Αυτές ορίζονται, σε αντίθεση με της διτονικές, ως οι εικόνες στις οποίες τα γειτονικά σημεία δεν διαφέρουν σημαντικά μεταξύ τους δηλαδή χαρακτηρίζονται από ομαλές τονικές διαβαθμίσεις. Υπάρχουν δύο είδη εικόνων συνεχούς τόνου: κλίμακας του γκριζου (gray scale) και έγχρωμες (colour). Το πρώτο είδος βρίσκει παρόμοιες εφαρμογές με τις διτονικές. Η διαφορά είναι ότι τα έγγραφα μπορούν τώρα να έχουν και εικόνες οι οποίες αποδίδονται με διαβαθμίσεις του γκριζου. Για παράδειγμα, ιατρικές φωτογραφίες αποτέλεσμα ακτινογραφιών ή υπερηχογραφήμάτων μπορούν να αποδοθούν ικανοποιητικά από εικόνες κλίμακας του γκριζου. Οι έγχρωμες εικόνες, όπως είναι φυσικό βρίσκουν τη μεγαλύτερη χρήση και έχουν το μεγαλύτερο ενδιαφέρον. Χρησιμοποιούνται τόσο σε επαγγελματικές όσο και σε εκπαιδευτικές και ψυχαγωγικές εφαρμογές. Αναμφισβήτητα, οι χρήση εικόνας βελτιώνει της υπάρχουσες εφαρμογές.

Σε συνδυασμό όμως και με άλλες τεχνολογίες, όπως η αναγνώριση προτύπων και τα έμπειρα συστήματα, ανοίγουν το δρόμο για εντελώς νέες εφαρμογές. Ένα παράδειγμα είναι η αυτόματη ταυτοποίηση ατόμων με βάση τα δακτυλικά αποτυπώματα ή κάποια φωτογραφία, εφαρμογή που είναι χρήσιμη σε συστήματα ασφαλείας.

## 2.2 Πρότυπο JPEG

Το JPEG είναι ένα πρότυπο του ISO το οποίο σχεδιάστηκε από την ομάδα Joint Photographic Expert Group σε συνεργασία με την Διεθνή Ένωση Τηλεπικοινωνιών (International Telecommunication Union, ITU-TS). Πρόκειται ίσως για το σημαντικότερο και πιο αποτελεσματικό πρότυπο συμπίεσης εικόνας, το οποίο κερδίζει συνεχώς έδαφος στις εφαρμογές πολυμέσων σε όλες τις πλατφόρμες[6]. Γι' αυτόν τον λόγο, θα σταθούμε στις γενικές αρχές του λίγο παραπάνω. Με δυο λόγια θα μπορούσαμε να περιγράψουμε το JPEG ως εξής:

Το JPEG είναι ένα πρότυπο συμπίεσης εικόνων συνεχούς τόνου, είτε έγχρωμων είτε κλίμακας του γκριζου. Χρησιμοποιεί ένα συνδυασμό τεχνικών διακριτού συνημιτονικού μετασχηματισμού Fourier, κβαντοποίησης, περιορισμού των επαναλαμβανόμενων χαρακτήρων και κωδικοποίησης Huffman και υποστηρίζει διάφορους τρόπους λειτουργίας. Μπορεί να έχει απώλειες με διάφορους συνδυασμούς λόγου συμπίεσης-ποιότητας ή και να λειτουργεί χωρίς απώλειες.

Το JPEG έχει τέσσερις ρυθμούς λειτουργίας:

- Διαδοχική κωδικοποίηση (sequential encoding)

Σε αυτόν το ρυθμό λειτουργίας το JPEG λειτουργεί με απώλειες και γίνεται μια μόνο σάρωση της εικόνας από αριστερά προς τα δεξιά και από πάνω προς τα κάτω. Πρόκειται για τον συνηθέστερο ρυθμό λειτουργίας.

- Προοδευτική κωδικοποίηση (progressive encoding)

Και αυτός ο ρυθμός λειτουργίας παρουσιάζει απώλειες. Η κωδικοποίηση γίνεται όμως με διαδοχικά περάσματα.

- Κωδικοποίηση χωρίς απώλειες (lossless encoding)

Αντίθετα με τις υπόλοιπες περιπτώσεις, το αποτέλεσμα αυτού του τρόπου συμπίεσης είναι πλήρως αντιστρέψιμο.

- Ιεραρχική κωδικοποίηση (hierarchical encoding)

Η κωδικοποίηση συνίσταται από διάφορα επίπεδα ευκρίνειας, τα οποία μπορούν να αποκωδικοποιηθούν ξεχωριστά.

### 2.2.1 Βασικά Πλεονεκτήματα Χρήσης JPEG

Η χρήση της μεθόδου συμπίεσης JPEG έχει τα εξής πλεονεκτήματα:

- Τα αρχεία των συμπιεσμένων εικόνων μπορούν να γίνουν πολύ μικρά σε μέγεθος και έτσι γίνεται γρήγορη η μετάδοση τους μέσα από δίκτυα, όπως επίσης γίνεται πραγματοποιήσιμη η περιληπτική αρχειοθέτηση σε βιβλιοθήκες εικόνων. Μια μέση συμπίεση μετατρέπει μια έγχρωμη εικόνα μεγέθους 2 MB σε ένα αρχείο 100 KB δηλαδή συμπίεση 20 προς 1, ενώ η αντίστοιχη μέθοδος GIF προσφέρει συμπίεση της τάξεως του 5 προς 1.
- Το JPEG αποθηκεύει την εικόνα σε 24 bits ανά pixel (σε 16 εκατομμύρια χρώματα), την στιγμή που η μέθοδος GIF μπορεί να αποθηκεύει μόνο σε 8 bits ανά pixel (σε 256 ή λιγότερα χρώματα. Έτσι με τις συνεχώς αναβαθμιζόμενες δυνατότητες σε hardware, οι υπολογιστές θα υιοθετήσουν ως νέο πρότυπο για προβολή εικόνων, αυτό που θα αξιοποιεί τις δυνατότητές τους αυτές στον μεγαλύτερο βαθμό[7].
- Τέλος, η ανταλλαγή των αρχείων JPEG μεταξύ χρηστών με διαφορετικό hardware είναι πιο εύκολη από την αντίστοιχη με GIF γιατί δεν προδικάζει πόσα χρώματα θα χρησιμοποιήσει (βλέπε κβαντοποίηση χρωμάτων) και έτσι τα αρχεία αυτά είναι περισσότερο κατάλληλα για χρήση μέσω World Wide Web.

Το αντίτιμο σε αυτά τα πλεονεκτήματα του JPEG είναι ότι χρειάζεται περισσότερος χρόνος για να γίνει η αποκωδικοποίηση και η προβολή σε σχέση με τις άλλες μεθόδους. Όταν όμως πρόκειται να παρεμβληθούν δικτυακές ή τηλεφωνικές μεταδόσεις, τότε το κέρδος που έχουμε στην σχέση χρόνου μετάδοσης αρχείου / χρόνου αποσυμπιεστής εικόνας είναι μεγάλο. Ένα άλλο πρόβλημα που μπορούσε να εμφανιστεί σε κάποιον χρήστη ήταν η έλλειψη υποστήριξης του προτύπου JPEG από τον viewer που διέθετε, οπότε έπρεπε να κάνει μετατροπή σε άλλο πρότυπο για την προβολή της εικόνας, γεγονός που ισοδυναμούσε με περισσότερο χρόνο στη χρήση του JPEG. Με την εξάπλωση όμως του προτύπου αυτού και την ύπαρξη τόσων JPEG viewers, σήμερα αυτό το πρόβλημα τείνει να εξαλειφθεί. Τέλος, ένα γενικότερο μειονέκτημα της συμπίεσης με απώλειες παρουσιάζεται όταν κάνουμε επανειλημμένες συμπίεσεις και αποσυμπιέσεις μίας εικόνας.

## 2.3 JPEG Χωρίς Απώλειες

Υπάρχει μεγάλη σύγχυση σ' αυτό το θέμα για την ύπαρξη μιας τέτοιας μορφής συμπίεσης JPEG[2]. Το στάνταρτ JPEG περιλαμβάνει έναν αλγόριθμο συμπίεσης που είναι πραγματικά χωρίς απώλειες, δηλαδή έναν αλγόριθμο που εξασφαλίζει ότι το αποτέλεσμα της αποσυμπίεσης του είναι ψηφίο προς ψηφίο ίδιο με την αρχική είσοδο. Ωστόσο, αυτός ο αλγόριθμος δεν έχει σχεδόν τίποτα κοινό με τον απλό (αλλά με απώλειες) JPEG αλγόριθμο, και προσφέρει πολύ μικρότερη συμπίεση.

Ο βαθμός συμπίεσης που προσφέρει το JPEG χωρίς απώλειες είναι 2:1 (τυπική τιμή). Δουλεύει καλά μόνο για εικόνες με συνεχή τόνο χρώματος ενώ σε εικόνες που χρησιμοποιούν χρώματα παλέτας και σε εικόνες με μικρό βάθος χρώματος, η συμπίεση που παρέχει δεν είναι χρήσιμη. (Τα νεότερα αποτελέσματα δείχνουν ότι το PNG ξεπερνά κατά πολύ το χωρίς απώλειες JPEG στις περισσότερες εικόνες. Προφανώς μόλις το PNG διαδοθεί περισσότερο, το χωρίς απώλειες JPEG θα καταργηθεί).

Υπάρχουν πολύ λίγες υλοποιήσεις του πραγματικά χωρίς απώλειες JPEG. Μια απ' αυτές είναι το RVPG.

Η μέθοδος αυτή λοιπόν, αποτελεί ανεξάρτητο λογισμικό από το βασικό JPEG, δεν χρησιμοποιεί ΔΣΜ (Διακριτός Συνημιτονικός Μετασχηματισμός), αφού η χρήση του ΔΣΜ λόγω των σφαλμάτων στρογγυλοποίησης δημιουργεί απώλειες. Για τον ίδιο λόγο δεν θα πρέπει να χρησιμοποιείται μετατροπή χρωματικής περιοχής ή δειγματοληψία, αν και αυτές οι λειτουργίες επιτρέπονται από το πρότυπο αυτό. Η μέθοδος συμπίεσης χωρίς απώλειες αναμενόμενης τιμής για το pixel αυτό. Η αναμενόμενη τιμή είναι μια συνάρτηση των ήδη γνωστών pixels που βρίσκονται πάνω και αριστερά από το παρόν και έχουν ήδη μεταδοθεί. Ως συνάρτηση μπορεί να χρησιμοποιηθεί ο μέσος όρος ή άλλες 8 συναρτήσεις που προσφέρονται από το ίδιο το πρόγραμμα. Η ακολουθία των διαφορών αυτών κωδικοποιείται με χρήση του ίδιου κώδικα που χρησιμοποιήσαμε στο JPEG με απώλειες (δηλαδή του Huffman ή τον αριθμητικό).

Αν ρυθμίσουμε ένα απλό JPEG στη μέγιστη δυνατή ποιότητα δε σημαίνει ότι θα αποθηκευτούν τα δεδομένα χωρίς απώλειες. Το JPEG χωρίς απώλειες είναι μια τελείως διαφορετική μέθοδος. Ακόμα και στη καλύτερη δυνατή ποιότητα, το απλό JPEG δεν μπορεί να είναι χωρίς απώλειες γιατί έχουμε λάθη από στρογγυλοποιήσεις στους διάφορους υπολογισμούς. Τα λάθη αυτά είναι σχεδόν πάντα πολύ μικρά για να γίνουν εμφανή, αλλά θα εμφανιστούν αν υποβάλλουμε την εικόνα σε πολλούς κύκλους συμπίεσης.

Το JPEG χωρίς απώλειες με την χρήση κώδικα Huffman δεν έχει τα καλύτερα αποτελέσματα που θα μπορούμε να περιμένουμε, ενώ ο αριθμητικός κώδικας είναι πιο ανταγωνιστικός αλλά και πάλι όχι η καλύτερη λύση για συμπίεση χωρίς απώλειες. Ο κυριότερος λόγος που υπάρχει η μέθοδος JPEG χωρίς απώλειες είναι γιατί στην ιεραρχική μέθοδο το τελευταίο πέρασμα μπορεί να είναι με κωδικοποίηση χωρίς απώλειες την υπολειπόμενων διαφορών ώστε να επιτευχθεί ολική ακρίβεια. Ακόμα όμως και τότε δεν είναι σίγουρο ότι δεν θα έχουμε απώλειες αφού πρέπει ο κωδικοποιητής και ο αποκωδικοποιητής να έχουν παρόμοιο σφάλμα στρωγγυλοποίησης.

Εξάλλου, πολλές υλοποιήσεις του JPEG δεν επιτρέπουν να χρησιμοποιηθεί με τη μέγιστη ρύθμιση ποιότητας, αφού θεωρείτε πολύ παράδοξο να χρησιμοποιείτε μ' αυτόν τον τρόπο το JPEG. Για παράδειγμα στο πρόγραμμα IJG JPEG, πρέπει να δηλώσουμε όχι μόνο «-quality 100» αλλά και «-sample 1x1» για να εξαλείψουμε το χάσιμο της πληροφορίας. Τα αρχεία που προκύπτουν είναι κατά πολύ μεγαλύτερα και λίγο καλύτερα σε ποιότητα από τα αρχεία που δημιουργούνται με λογικότερες ρυθμίσεις.

Εκτός αυτού εξακολουθούν να έχουν κάποιες απώλειες! Αν πραγματικά χρειαζόμαστε μία αποθήκευση χωρίς καθόλου απώλειες, δεν προσπαθούμε ποτέ να την πετύχουμε με χρήση του απλού JPEG.

## 2.4 Πρότυπο JBIG

Το JBIG είναι ένα πρότυπο ISO. Αυτό το πρότυπο καθορίζει έναν αλγόριθμο συμπίεσης χωρίς απώλειες (lossless) για δυαδικές (binary) εικόνες. Ο στόχος του JBIG είναι να αντικαταστήσει τους τωρινούς αλλά λιγότερο αποτελεσματικούς αλγορίθμους της ομάδας 3 και 4 των fax αλγορίθμων. Στις binary εικόνες, κάθε pixel αναπαρίσταται από ένα bit, για παράδειγμα μια εικόνα αντιστοιχεί σε ένα σχέδιο από bit. Το πρότυπο συμπίεσης JBIG βασίζεται σε ένα συνδυασμό προβλέψιμης και αριθμητικής κωδικοποίησης. Το πρότυπο JBIG[5] μπορεί επίσης να χρησιμοποιηθεί σε ασπρόμαυρες σε ασπρόμαυρες ή και σε έγχρωμες φωτογραφίες για lossless συμπίεση χρησιμοποιώντας απλά τον αλγόριθμο του 1 bit σχεδίου κάθε φορά. Αυτό το πρότυπο δουλεύει αρκετά καλά μέχρι και για 6 bits ανά pixel, από κει και πάνω το πρότυπο JPEG δουλεύει καλύτερα.

## 2.5 Πρότυπο GIF

Gif είναι το ακρωνύμιο του Graphics Interchange Format. Το GIF είναι μια δημιουργία της CompuServe Inc. και χρησιμοποιείται για αποθήκευση πολλαπλών raster εικόνων σε ένα μόνο αρχείο για μεταφορά μεταξύ πλατφόρμων και συστημάτων. Το GIF είναι ένα format συναλλαγής κι αποθήκευσης που υποστηρίζεται από πολλές εφαρμογές. Πολλές από τις εικόνες που αποθηκεύονται σ' αυτό είναι υψηλής - ποιότητας εικόνες ανθρώπων, τοπίων, αστρο-φωτογραφίες κ.α.

Υπάρχουν δύο εκδόσεις του προσδιορισμού GIF[3], και οι δύο είναι διαδεδομένες. Η αυθεντική έκδοση ήταν η GIF87a που παρουσιάστηκε το Μάιο του 1987, και δημιουργήθηκαν πολλές εικόνες σ' αυτό το format. Η πιο νέα έκδοση η GIF89a, που παρουσιάστηκε τον Ιούλιο του 1989, προσθέτει αρκετές δυνατότητες συμπεριλαμβανομένης και της ικανότητας να αποθηκεύει δεδομένα κειμένου και γραφικών μέσα στο ίδιο αρχείο. Σε μια εφαρμογή, πάντως θα ήταν χρήσιμη η υποστήριξη και την δυο εκδόσεων του GIF.

Το GIF format μπορεί να αποθηκεύει raster δεδομένα με βάθος pixel από 1 έως 8 bits. Ο μεγαλύτερος όγκος των GIF αρχείων περιέχει 16 ή 256 χρώματα που πλησιάζουν την ποιότητα των φωτογραφιών. Εικόνες gray-scale όπως αυτές που παράγονται από σαρωτές αποθηκεύονται στο GIF, ωστόσο, μονόχρωμα γραφικά, αποθηκεύονται σπάνια στο GIF. Οι εικόνες αποθηκεύονται πάντα χρησιμοποιώντας το RGB μοντέλο χρώματος και χρησιμοποιώντας δεδομένα παλέτας.

Παρόλο που ο όγκος των GIF αρχείων βρίσκονται στο περιβάλλον MS-DOS της Intel, το GIF ούτε σχετίζεται, ούτε δημιουργήθηκε για τις ανάγκες κάποιου συγκεκριμένου πακέτου ή υπερβάλλοντος. Ωστόσο, τα περισσότερα πακέτα, και διαβάζουν και αποθηκεύουν δεδομένα εικόνας GIF. Τέτοια προγράμματα τα σχεδιαστικά, ζωγραφικής, λογισμικό σάρωσης και video και τα περισσότερα προγράμματα εμφάνισης αρχείων εικόνας και μετατροπής αρχείων. Το GIF ήταν προορισμένο να επιτρέπει την εύκολη διακίνηση και προβολή των δεδομένων εικόνας που αποθηκεύονται σε τοπικά ή απομακρυσμένα συστήματα υπολογιστών. Άλλωστε, πολύ σημαντικό σημείο για το GIF format, είναι η αποτελεσματική μετάδοση των δεδομένων μέσω τηλεφωνικών γραμμών.

### 2.5.1 Προβλήματα GIF

Το GIF είναι ένα format που έχει οριστεί και τεκμηριωθεί σωστά, σε ευρεία χρήση, το οποίο διαβάζεται εύκολα και γρήγορα και αποσυμπίεζεται εύκολα. Παρόλα αυτά έχουν παρατηρηθεί και κάποια προβλήματα[1].

Προβλήματα έχουν παρατηρηθεί όταν εφαρμογές που διαβάζουν μόνο GIF87a αρχεία εικόνας, προσπαθούν να προβάλουν GIF89a αρχεία εικόνας. Παρόλο που το GIF89a format μοιάζει πολύ με το GIF87a, περιέχει αρκετά πρόσθετα μπλοκ πληροφορίας που δεν ορίζονται στο 87ζ προσδιορισμό. Γι' αυτό τον λόγο τα GIF89a αρχεία εικόνων δεν μπορούν να διαβαστούν και να εμφανιστούν σωστά από αυτές τις εφαρμογές. Ο λόγος που πολλά από αυτά τα προγράμματα δεν μπορούν να εμφανίσουν ένα GIF89a αρχείο εικόνας, είναι επειδή δεν αναγνωρίζουν τον αριθμό έκδοσης που δηλώνεται στην επικεφαλίδα κάθε GIF αρχείου, αλλάζοντας όμως αυτόν τον αριθμό από «89a» σε «87a» το πρόγραμμα θα μπορέσει να εμφανίσει τα δεδομένα εικόνας αλλά όχι σωστά.

Το GIF είναι ακόμα ικανό να αποθηκεύει πολλαπλές εικόνες ανά αρχείο αλλά το μεγαλύτερο μέρος των GIF αρχείων περιέχουν μόνο μία εικόνα. Άλλωστε λίγες είναι οι εφαρμογές εμφάνισης που υποστηρίζουν τη προβολή GIF αρχείων με πολλαπλές εικόνες, σε αντίθεση με τον μεγαλύτερο όγκο των εφαρμογών, οι οποίες συνήθως εμφανίζουν μόνο την πρώτη κατά σειρά εικόνα του GIF αρχείου και τις υπόλοιπες τις αγνοούν.

### 2.5.2 Οργάνωση του Αρχείου

Το GIF αποτελείται από μία σειρά πακέτων δεδομένων, που λέγονται *μπλοκ (block)*, μαζί με πρόσθετες πληροφορίες[4]. Λόγω της οργάνωσης τους, τα GIF αρχεία πρέπει να διαβάζονται σαν να ήταν μία συνεχή ροή δεδομένων. Τα διάφορα μπλοκ και υπομπλοκ δεδομένων που ορίζονται από το GIF μπορεί να βρεθούν σχεδόν παντού μέσα στο αρχείο.

Υπάρχει ένας αριθμός διαφορετικών κατηγοριών μπλοκ δεδομένων, και καθένα από τα διάφορα οριζόμενα μπλοκ εμπίπτει σε μια από τις κατηγορίες. Τα μπλοκ εκτός από την αποθήκευση πεδίων πληροφορίας μπορούν να περιέχουν υπό-μπλοκ. Κάθε υπομπλοκ δεδομένων ξεκινάει με ένα byte μέτρησης. Η τιμή αυτού του μετρητή μπορεί να κυμαίνεται από 1-225 και δηλώνει τον αριθμό των bytes των δεδομένων στο υπομπλοκ. Τα μπλοκ δεδομένων ακολουθούν αμέσως μετά το byte μέτρησης. Μία συνεχής ομάδα από μπλοκ δεδομένων τερματίζει με ένα byte μηδενικής τιμής, που δηλώνει ότι δεν ακολουθούν άλλα δεδομένα.

### 2.5.3 GIF 87a

Η έκδοση GIF87a είναι το αυθεντικό GIF format που παρουσιάστηκε τον Μάιο του 1987 και διαβάζεται από το μεγαλύτερο μέρος των εφαρμογών που υποστηρίζουν το GIF format[5].



Κάθε αρχείο πάντα αρχίζει με μια Επικεφαλίδα (Header) και ένα Logical Screen Descriptor. Ένας Global μπορεί να εμφανιστεί κάποιες φορές μετά το Logical Screen Descriptor. Κάθε ένα από αυτά τα τρία τμήματα βρίσκεται πάντα στην ίδια μετατόπιση από την αρχή του αρχείου. Κάθε εικόνα που αποθηκεύεται σ' ένα αρχείο περιέχει έναν Logical Image Descriptor, έναν προαιρετικό Local Color Table και ένα μπλοκ δεδομένων εικόνας. Το τελευταίο πεδίο σε κάθε GIF αρχείο είναι ένας χαρακτήρας τερματισμού ο οποίος δηλώνει το τέλος της GIF ροής δεδομένων.

## 2.6 Fractal Image Compression (Απειροστική Συμπίεση Εικόνας)

### 2.6.1 Γενικά στοιχεία σχετικά με την απειροστική συμπίεση εικόνας

1. Είναι μια καινούρια και πολλά υποσχόμενη τεχνολογία. Το αν είναι ανώτερη της συμπίεσης JPEG είναι ανοιχτό σε διάλογο.
2. Είναι μέθοδος συμπίεσης χωρίς απώλειες.
3. Τα απειροστικά τμήματα στην απειροστική συμπίεση εικόνας είναι επαναλαμβανόμενα συστήματα συναρτήσεων (Iterated Function Systems).
4. Είναι μια μορφή κβαντικού διανύσματος, κάτι που χρειάζεται ειδική ορολογία για να καταλάβουμε.
5. Η αύξηση της ανάλυσης είναι ένα πανίσχυρο σημείο της μεθόδου αλλά δεν είναι κάτι το μαγικό που πετυχαίνει συμπίεση 1000:1.
6. Η συμπίεση είναι αρχή, η αποσυμπίεση είναι γρήγορη.
7. Η τεχνολογία είναι πατενταρισμένη.

### 2.6.2 Λειτουργία απειροστικής συμπίεση εικόνας

Τα απειροστικά τμήματα που χρησιμοποιούνται στην απειροστική συμπίεση της εικόνας δεν είναι αυτά που γνωρίζουμε από την ανάλυση, αλλά από τη θεωρία της επαναλαμβανόμενης συνάρτησης[1]. Ο μαθηματικός Heinz-Otto Peitgen μας δίνει μια ωραία εισαγωγή στη θεωρία αυτή χρησιμοποιώντας μεταφορικά ένα αντιγραφικό μηχάνημα πολλαπλής ελάττωσης (Multiple Reduction Copying Machine).

Το MRCM υποτίθεται πως είναι ένα απλό αντιγραφικό μηχάνημα, με τις εξής διαφορές:

Υπάρχουν πολλές ρυθμίσεις των φακών που μπορούν να δημιουργήσουν πολλές επικαλυπτόμενες κόπιες του αρχικού.

Κάθε μία ρύθμιση των φακών μειώνει το μέγεθος του αρχικού.

Το αντιγραφικό λειτουργεί με βρόγχο ανάδρασης, με την έξοδο της κάθε βαθμίδας να είναι είσοδος στην επόμενη. Η αρχική είσοδος μπορεί να είναι οτιδήποτε.

Το πρώτο είναι αυτό που κάνει το IFS να είναι σύστημα. Το τρίτο είναι αυτό που το κάνει επαναλαμβανόμενο. Όσο για το δεύτερο, υπονοεί ότι οι συναρτήσεις μιας επαναλαμβανόμενης συνάρτησης είναι συστολικές.

Το IFS λοιπόν είναι μια ομάδα συστολικών μετασχηματισμών που σχεδιάζουν από ένα συγκεκριμένο ορθογώνιο φυσικού σχεδίου σε μικρότερα τμήματα αυτού του ορθογωνίου.

Αυτό το IFS αποτελείται από τρεις συνθετικούς μετασχηματισμούς (τρεις διαφορετικοί φακοί στο μηχάνημα MRCM). Κάθε ένας απ' αυτούς συρρικνώνει το αρχικό κατά 2, και μετά μεταφράζει το αποτέλεσμα σε μια νέα τοποθεσία. Μπορεί προαιρετικά να ρυθμίσει και να μετακινήσει τη φωτεινότητα του ορθογωνίου, με ένα τρόπο παρόμοιο με αυτό της ρύθμισης που κάνουμε στη φωτεινότητα και τη διαύγεια της τηλεόρασης.

### 2.6.3 Απώλεια κλίμακας και αύξηση της ανάλυσης

Όταν μία συσκευή, όπως η φωτογραφική μηχανή, παίρνει μία εικόνα, χρησιμοποιεί μία κλίμακα που καθορίζεται από τη στοιχειώδη ανάλυση αυτής της συσκευής. Αν χρησιμοποιήσουμε κάποιο πρόγραμμα για να μεγεθύνουμε αυτή την εικόνα, από ένα σημείο και μετά δεν βλέπουμε κάτι παραπάνω σε λεπτομέρεια, παρά μόνο μεγαλύτερα εικονοστοιχεία [3].

Μία απειροστική εικόνα είναι διαφορετική. Με κάθε επανάληψη, η λεπτομέρεια που δημιουργείται είναι έξοχη σε μεγαλύτερη ανάλυση. Οριακά, λεπτομέρεια που αντιστοιχεί με ακρίβεια στην πραγματικότητα, δημιουργείται σε όλα τα επίπεδα ανάλυσης, μέχρι και το άπειρο. Επειδή δεν υπάρχει όριο στο οποίο να σταματούν οι απειροστικές εικόνες, θεωρούμε ότι δεν έχουν κλίμακα.

Πρακτικά, αυτό σημαίνει ότι αν μεγεθύνουμε μία απειροστική εικόνα, θα εξακολουθεί να δείχνει “όπως πρέπει” χωρίς τα φαινόμενα που εμφανίζει η μεγέθυνση των εικονοστοιχείων. Η σημαντικότητα των παραπάνω έχει δημιουργήσει μερικές διαφωνίες και γι αυτό θα πρέπει να πούμε τα παρακάτω:

Η Iterated Systems υποστηρίζει τα εξής: Πάρτε ένα πορτραίτο που είναι, ας πούμε, μια ασπρόμαυρη εικόνα κι έχει μέγεθος 250x250 pixel, με ένα byte ανά pixel. Την περνάτε από το δικό της πρόγραμμα κι έχετε ένα αρχείο μεγέθους 2500 byte (συντελεστής συμπίεσης = 25:1). Μεγεθύνετε τώρα τα μαλλιά του ατόμου της φωτογραφίας τέσσερις φορές. Τι βλέπετε; Μια εικόνα που μοιάζει πάλι με μαλλιά. Επομένως, είναι σα να έχετε μια εικόνα μεγέθους 1000x1000. Άρα και ο συνολικός συντελεστής συμπίεσης που καταφέρατε να πάρετε είναι  $25 \times 16 = 400$ .

Αλλά τα πράγματα δεν είναι ακριβώς έτσι. Η λεπτομέρεια που παίρνουμε δεν διατηρήθηκε από την αρχική, αλλά δημιουργήθηκε. Με λίγη τύχη μοιάζει με αυτό που θα έπρεπε να μοιάζει αλλά μην είστε και πολύ σίγουρα. Αν για παράδειγμα μεγεθύνουμε το πρόσωπο, δε θα δούμε τους πόρους του.

Αντικειμενικά, αυτό που προσφέρει η απειροστική συμπίεση εικόνας είναι μια αναπτυγμένη μορφή παρεμβολής. Αυτό είναι κάτι χρήσιμο κι ελκυστικό. Χρήσιμο για παράδειγμα, σε όσους ασχολούνται με τη δημιουργία γραφικών ή σ' εκείνους που θέλουν να τοπώσουν σε μια συσκευή υψηλής ανάλυσης. Όμως δεν προσφέρει εξωπραγματικούς συντελεστές συμπίεσης.

Με βάση τα παραπάνω, πρέπει να απαντήσουμε στο ερώτημα τι εννοούμε με τον όρο αύξηση της ανάλυσης. Πρόκειται για τη διαδικασία που συμπεζούμε μια εικόνα, την μεγεθύνουμε σε μια μεγαλύτερη ανάλυση και τη σώνουμε, άσχετα με τον αν χρησιμοποιούμε το IFS. Με άλλα λόγια, η συμπεσμένη απειροστικά εικόνα είναι κάποιο μέσο για να πετύχουμε το σκοπό μας και όχι ο ίδιος ο σκοπός.

## 2.7 Απειροστικότητα - JPEG

Η μεγάλη ειρωνεία για την επιτροπή κωδικοποίησης είναι ότι μεγάλη φροντίδα δίνεται προκειμένου να μετρηθεί ακριβώς και να προσδιορισθεί το ποσοστό λάθους σε μία συμπεσμένη εικόνα, και γίνεται μεγάλη προσπάθεια προκειμένου να μειωθεί το λάθος που μετρείται το οποίο πολύ συχνά είναι αμφίβολο[4]. Αυτές οι μετρήσεις περιλαμβάνουν λόγο σήματος προς θόρυβο, μέθοδο ελάχιστων τετραγώνων και απόλυτη τιμή του λάθους. Ένα απλό παράδειγμα είναι η συνεχής μετακίνηση (shift): προσθέτουμε την τιμή 10 σε κάθε pixel. Σταθερές μετρήσεις λάθους υποδεικνύουν μεγάλη παραμόρφωση, αλλά η εικόνα έχει γίνει ελάχιστα πιο φωτεινή.

Με σεβασμό σε αυτές τις αμφίβολες μετρήσεις λάθους, και με το φόβο της υπεραπλούστευσης, τα αποτελέσματα αποκαλύπτουν τα ακόλουθα:

Για μικρούς συντελεστές συμπίεσης το JPEG είναι καλύτερο, για υψηλούς συντελεστές τα η απειροστική κωδικοποίηση είναι καλύτερη. Το διαχωριστικό σημείο διαφέρει, αλλά συνήθως είναι γύρω στο 40:1. Η τιμή αυτή προμηνύει ανωτερότητα του JPEG αφού έτσι κι αλλιώς πέρα από αυτό το βαθμό συμπίεσης οι εικόνες έχουν συνήθως τέτοια αταξία που δεν αξίζει να χρησιμοποιηθεί.

Υποστηρικτές της απειροστικής συμπίεσης υποστηρίζουν ότι ο λόγος σήματος προς θόρυβο δεν είναι καλό μέτρο για τη μέτρηση του λάθους και ότι η παραμορφωμένη εικόνα που δίνει είναι πολύ πιο “φυσική”, από τα τετράγωνα μπλοκ του JPEG, τόσο στις χαμηλές όσο και στις υψηλές αναλύσεις. Αυτό είναι ένα λογικό επιχείρημα αλλά σε καμία περίπτωση δεν είναι αποδεκτό παγκόσμια. Ίσως αυτό που χρειάζεται η επιτροπή κωδικοποίησης να είναι ένας τρόπος που θα μετρά με ακρίβεια και αντικειμενικότητα την εντύπωση που δίνει το σήμα (εδώ η εικόνα) στον άνθρωπο. Μέχρι τότε τα μάτια σας είναι οι καλύτεροι κριτές.

## 2.8 JPEG - GIF

Σε γενικές γραμμές, η μέθοδος συμπίεσης JPEG είναι ανώτερη από την αντίστοιχη GIF, όταν δουλεύουμε με έγχρωμες ή ασπρόμαυρες «ρεαλιστικές» σκηνές, δηλαδή φωτογραφίες από scanner και παρόμοιες εικόνες. Συνεχείς αλλαγές σε χρώμα[2], όπως αυτές που παρουσιάζονται σε σκιασμένες ή φωτισμένες περιοχές, παρουσιάζονται πιο πιστά και καταλαμβάνοντας λιγότερο χώρο θα αρχείο JPEG και όχι GIF.

Παρόλα αυτά η μέθοδος GIF δεν πρόκειται να εγκαταλειφθεί αφού κάνει πολύ καλύτερη συμπίεση με εικόνες που έχουν λίγα χρώματα, όπως απλά ή γραμμικά σχέδια.

Σε αυτές τις εικόνες, η GIF συμπίεση είναι εντελώς χωρίς απώλειες και επιπλέον κάνει συχνά μεγαλύτερη συμπίεση από την JPEG, γιατί μεγάλες περιοχές στοιχείων της εικόνας (pixels), που είναι ακριβώς το ίδιο χρώμα, συμπιέζονται πολύ καλύτερα με την GIF. Για παράδειγμα, ένα μονόχρωμο περιθώριο με συμπίεση GIF, έχει σχετικά μικρότερο κόστος σε αποθηκεύσιμο χώρο από την JPEG και είναι καλύτερο να αποφεύγεται στην δεύτερη περίπτωση.

Όσον αφορά τις εικόνες που έχουν δημιουργηθεί με την χρήση υπολογιστή, όπως σκηνές από ανίχνευση ακτίνας (ray-tracing), βρίσκονται ανάμεσα στις φωτογραφίες και στα απλά σχέδια από άποψη πολυπλοκότητας και όσο πιο πολύπλοκη είναι η εικόνα, τόσο πιο πιθανό είναι να έχει καλύτερα αποτελέσματα η JPEG συμπίεση πάνω της.

Η συμπίεση JPEG παρουσιάζει προβλήματα όταν στην εικόνα υπάρχουν οξείες ακμές όπως μια σειρά μαύρων pixels δίπλα σε μια σειρά λευκών pixels. Σε αυτές τις περιπτώσεις οι ακμές παρουσιάζονται θολές εκτός αν χρησιμοποιήσουμε πολύ υψηλή ρύθμιση ποιότητας, ενώ τέτοιες ακμές δεν συναντιούνται σε φωτογραφίες από scanner, αλλά είναι πολύ συχνές σε αρχεία GIF (ευθείες γραμμές κλπ.).

Επίσης ένα αντίστοιχο πρόβλημα θολούρας παρουσιάζεται όταν έχουμε κείμενο με πολύ μικρό ύψος γραμμάτων, οπότε εάν έχουμε ένα τέτοιο αρχείο με συμπίεση GIF, θα ήταν καλύτερα να μην το μετατρέψουμε σε JPEG αρχείο. Εάν, τέλος θέλουμε να προσθέσουμε κείμενο με μικρό ύψος χαρακτήρων σε μια JPEG εικόνα, θα ήταν καλύτερα να το περάσουμε ως σχόλιο (comment) σε αυτή, αν και υπάρχει ο κίνδυνος μερικοί viewers να αγνοήσουν το κείμενο-σχόλιο κατά την προβολή της εικόνας. Για τις καθαρά ασπρόμαυρες (2 χρώματα) εικόνες δεν πρέπει ποτέ να χρησιμοποιείται η συμπίεση JPEG, αφού αυτή δίνει χρήσιμα αποτελέσματα από 16 και πάνω αποχρώσεις του γκριζου. Σημειώνεται όμως ότι η GIF δεν έχει καθόλου απώλειες μέχρι τις 256 αποχρώσεις του γκριζου, ενώ η JPEG έχει και μάλιστα σημαντικές.

### 3. Ήχος

#### 3.1 Ήχος και Εφαρμογές

Μέχρι σήμερα, η βιομηχανία των υπολογιστών δεν έχει επιδείξει ιδιαίτερα μεγάλο ενδιαφέρον για τον ήχο στις εφαρμογές πολυμέσων. Αυτό οφείλεται κυρίως στην εστίαση των προσπαθειών στην εισαγωγή του video αλλά και στη δυσκολία καθορισμού της χρησιμότητας του ήχου σε επαγγελματικές εφαρμογές (business applications)[3]. Σαν αποτέλεσμα, ο ήχος στα πολυμέσα περιορίζεται στις ψυχαγωγικές και εκπαιδευτικές εφαρμογές.

Παρ' όλα αυτά, είναι φανερό ότι ο ήχος, κατάλληλα συνδυασμένος με τα άλλα είδη πληροφορίας, μπορεί να κάνει μια εφαρμογή πολυμέσων πιο αποτελεσματική. Ιδιαίτερα σε εκπαιδευτικές εφαρμογές και περίπτερα πληροφοριών (information kiosks) η αφήγηση και ο σχολιασμός των όσων παρουσιάζονται στην οθόνη βοηθά στην μετάδοση του μηνύματος ενώ η κατάλληλη ηχητικά υπόκρουση καθιστά την παρακολούθηση της εφαρμογής πιο ευχάριστη. Το μοναδικό χαρακτηριστικό του ήχου να γίνεται αντιληπτός χωρίς να έχουμε την προσοχή μας εστιασμένη καθιστά τα ηχητικά σήματα αναντικατάστατα στην προσοχής του χρήστη.

Υπάρχουν ορισμένες κατηγορίες εφαρμογών όπου ο ήχος αποτελεί την καρδιά του συστήματος. Εφαρμογές που έχουν ως αντικείμενο την μουσική ή ακόμα εφαρμογές που προορίζονται για ανθρώπους με προβλήματα όρασης κάνουν εκτενή και αποτελεσματική χρήση του ήχου.

Με την πρόοδο της τεχνολογίας, το ενδιαφέρον για την εφαρμογή της αναγνώρισης και σύνθεσης ομιλίας σε επαγγελματικές εφαρμογές μεγαλώνει. Ήδη έχουν εμφανιστεί τα πρώτα δείγματα συστημάτων χειρισμού ενός υπολογιστή με προφορικές εντολές και υπαγόρευσης κειμένου στον υπολογιστή.

### 3.2 Συμπύεση

#### Μουσική και υπολογιστές

Κάθε ήχος μπορεί να αποθηκευτεί στον υπολογιστή ως ψηφιοποιημένο ηχητικό σήμα. Αυτό το σήμα μπορεί να είναι συμπιεσμένο ή ασυμπιεστο. Σε κάθε όμως περίπτωση δεν περιέχει καμία σημασιολογική πληροφορία για τον ήχο που περιγράφει. Αυτή η μορφή αναπαράστασης του ήχου είναι το αντίστοιχο της χαρτογραφικής απεικόνισης των εικόνων. Σε αναλογία με τα γραφικά υπάρχει για τη μουσική το πρότυπο MIDI (Musical Instrument Digital Interface)[2].

Το πρότυπο αυτό αναπτύχθηκε στη αρχή της δεκαετίας του 80. Το MIDI καθορίζει πως κωδικοποιούνται τα διάφορα στοιχεία μιας μουσικής παρτιτούρας καθώς και τα όργανα που συμμετέχουν. Υπάρχει η δυνατότητα χρησιμοποίησης 127 οργάνων και ηχητικών εφέ. Το MIDI περιέχει και πρότυπα για την επικοινωνία μουσικών οργάνων με υπολογιστή. Ένας υπολογιστής με MIDI interface μπορεί να χειριστεί συσκευές που ακολουθούν αυτό το πρότυπο όπως ηλεκτρονικά synthesizers. Στις πιο πολλές κάρτες ήχου που προσφέρουν MIDI, η σύνθεση των ήχων των οργάνων γίνεται συνήθως με FM σύνθεση που δεν δίνει καλά αποτελέσματα. Σε πολλές όμως περιπτώσεις, περιέχουν αποθηκευμένα σε μνήμη ROM δείγματα πραγματικών οργάνων με αποτέλεσμα η μουσική MIDI να μοιάζει αρκετά με πραγματική.

Τα πλεονεκτήματα και μειονεκτήματα του και μειονεκτήματα του MIDI έναντι της ψηφιοποιημένης μουσικής είναι ανάλογα με αυτά των εικόνων έναντι των γραφικών. Υπάρχει μεγαλύτερη ευελιξία στην επεξεργασία της μουσικής MIDI ενώ απαιτείται και σημαντικά μικρότερος αποθηκευτικός χώρος. Από την άλλη πλευρά, υπάρχει υπολογιστικό κόστος για να μετατραπεί η συμβολική αναπαράσταση MIDI σε ακουστή κυματομορφή. Επίσης, το αποτέλεσμα εξαρτάται από τη συσκευή εξόδου και συνήθως είναι υποδεέστερο της ψηφιοποιημένης μουσικής.

Ο ήχος είναι γενικά δύσκολο να συμπιεστεί γιατί η ακοή είναι πιο ευαίσθητη στις αλλοιώσεις του ήχου σε σχέση με την όραση. Εξαιτίας του ενδιαφέροντος για συμπύεση του ήχου στην τηλεφωνία, έχουν αναπτυχθεί πολλές τεχνικές αποτελεσματικής κωδικοποίησης της ομιλίας. Στην μουσική τα πράγματα είναι πιο δύσκολα γιατί και οι απαιτήσεις ποιότητας είναι μεγαλύτερες και το εύρος ζώνης της είναι σημαντικά μεγαλύτερο από αυτό της ανθρώπινης φωνής.

Στον παρακάτω πίνακα παρουσιάζονται οι συχνότητες δειγματοληψίας και το μήκος της λέξης κατά την κβαντοποίηση για διάφορα πρότυπα συμπύεσης ήχου. Εκτός από την κατάλληλη επιλογή αυτών των παραμέτρων ανάλογα με τις απαιτήσεις της εφαρμογής, περαιτέρω συμπύεση μπορεί να επιτευχθεί με αφαίρεση των σιωπηλών τμημάτων και με καλύτερες μεθόδους κωδικοποίησης όπως: μη γραμμική PCM όπως λογαριθμική ή μ-law, διαφορική PCM και προσαρμοστική διαφορική PCM.

Συχνότητα Δειγματοληψίας (kHz)	Κβαντοποίηση (bits)	Τεχνική Κωδικοποίησης	Ποιότητα
44.1	16	PCM	Hi-fi
37.8	8	ADPCM	Hi-fi
37.8	8	ADPCM	FM μετάδοση (μουσική)
18.9		ADPCM	AM μετάδοση (ομιλία)
8	8	PCM	Τηλεφωνική

Σχήμα 1: Ηχητική ποιότητα και μέθοδος ψηφιοποίησης

### 3.3 Πρότυπα ITU

#### 3.3.1 Οπτικοακουστική Τηλεφωνία (Audiovisual Telephony)

Η οικογένεια προτύπων H.320 που δημοσιεύτηκαν από την ITU-TS και καλύπτουν εφαρμογές teleconferencing και videophony αποτελούν τη βάση επικοινωνίας μεταξύ συσκευών διαφορετικών κατασκευαστών κατασκευαστών. Το πρότυπο υποστηρίζει τόσο τηλεπικοινωνιακά δίκτυα υπολογιστών.[8] Στην οικογένεια αυτή περιλαμβάνονται και τη σειρά προτύπων ITU-G για συμπίεση ήχου. Τα χαρακτηριστικά αυτών των προτύπων φαίνονται στον επόμενο πίνακα.

Πρότυπο	Περιγραφή
G.711	PCM συχνοτήτων φωνής
G.722	ADPCM με συχνότητα 7kHz και απαιτούμενο εύρος ζώνης 64Kbits/s
G.723	ADPCM το οποίο λειτουργεί στα 24 Kbps
G.728	CELP κωδικοποίηση που απαιτεί 16Kbits/s

Σχήμα 2: Πρότυπα για κωδικοποίηση ήχου

- **G.721:**

Η σύσταση G.721 της ITU μετασχηματίζει ένα 64 Kbps bit stream σε ένα 32 Kbps stream. Το G.721 βασίζεται σε μια ADPCM τεχνική. Κάθε διαφορά τιμής κωδικοποιείται με 4 bits. Όπως και στο G.711 ο δειγματοληψίας είναι 8kHz.

- **G.722:**

Ο αντικειμενικός σκοπός του προτύπου G.722 είναι να εξασφαλίσει μια καλύτερη ποιότητα ήχου από το συμβατικό G.711 PCM πρότυπο ή την G.721 τεχνική συμπίεσης. Το G.722 βασίζεται σε μια sub-band ADPCM μέθοδο (SB-ADPCM). Για την μέθοδο αυτή αρκεί να πούμε ότι το διαθέσιμο εύρος ζώνης (bandwidth) χωρίζεται σε δύο ξεχωριστά sub-band στα οποία έχει εφαρμοστεί ο προσαρμόσιμος διαφορικός αλγόριθμος PCM. Το εύρος ζώνης ενός G.722 συμπιεσμένου σήματος κυμαίνεται μεταξύ 50 Hz και 7 kHz, δεδομένου ότι αυτό των G.711 σημάτων περιορίζεται σε 3.4 kHz. Το προκύπτων bit rate είναι 48,56 ή Kbps- στην τυπική λειτουργία, ο ρυθμός δειγματοληψίας είναι 16 kHz και το amplitude depth είναι 14 bits.

- **G.723:**

Το G.723 είναι ένα ακόμα lossy πρότυπο συμπίεσης βασισμένο στο ADPCM το οποίο λειτουργεί στα 24 Kbps. Η προκύπτουσα ποιότητα ήχου είναι κατώτερη από αυτήν του μη-συμπιεσμένου G.711 PCM προτύπου και αυτήν του προτύπου G.722 βασιζόμενο στο SB-ADPCM. Το G.723 χρησιμοποιείται σπάνια.

- **G.728:**

Το G.728 είναι ένα πρότυπο που στοχεύει σε χαμηλό bit rate. Λειτουργεί σε 16 Kbps αλλά το εύρος ζώνης περιορίζεται στα 3.4 kHz. Η προκύπτουσα ποιότητα ήχου είναι σημαντικά κατώτερη από αυτήν του G.711 και του G.722. Βασίζεται σε ένα σχέδιο κβαντοποίησης διανύσματος ονομαζόμενο low-delay code excited linear prediction (LD-CELP).

### 3.4 Τα πρότυπα MPEG

Η ομάδα Moving Pictures Experts Group (MPEG) του ISO/IEC αναπτύσσει πρότυπα για video και ήχο από το 1988[6]. Το πρότυπο που προέκυψε από την πρώτη φάση είναι το MPEG-1 ενώ η έρευνα συνεχίζεται στο MPEG-2. Για κάποιο χρονικό διάστημα υπήρχε σε εξέλιξη το MPEG-3 αλλά εγκαταλείφθηκε και ένα τμήμα του ενσωματώθηκε στο MPEG-2.

Κάθε ένα από αυτά στοχεύει σε εφαρμογές με διαφορετικές απαιτήσεις ποιότητας και εύρους ζώνης και εκτός από τεχνικές συμπίεσης κινούμενης εικόνας περιέχει και ένα κομμάτι που αναφέρεται στην συμπίεση του ήχου. Οι τεχνικές συμπίεσης ήχου δεν περιορίζονται στη φωνή αλλά είναι σχεδιασμένες για τη γενικότερη περίπτωση της μουσικής.

Στη συνέχεια θα περιγράψουμε σε συντομία τα βασικά χαρακτηριστικά του ηχητικού κομματιού του MPEG-1 ενώ στην επόμενη ενότητα θα επεκταθούμε και στην κινούμενη εικόνα.



- MPEG-1

Με δυο λόγια θα μπορούσαμε να περιγράψουμε το MPEG-1 ως εξής:

Το ηχητικό κομμάτι του MPEG-1 δεν αποτελεί έναν αλγόριθμο συμπίεσης, αλλά μια οικογένεια τριών διαφορετικών τεχνικών κωδικοποίησης και συμπίεσης ήχου. Αυτές οι οικογένειες ονομάζονται MPEG-Audio Layer-1, Layer-2, Layer-3.

Και οι τρεις αυτές τεχνικές στηρίζονται στην ίδιο αρχή: η συμπίεση γίνεται συνδυάζοντας ένα είδος κωδικοποίησης μετασχηματισμού και sub-band division. Οι διαφορές αυτών των τριών στρώματων εντοπίζονται στο τελικό στάδιο της κβαντοποίησης.

Τα βασικά βήματα που ακολουθούνται είναι:

- Χωρισμός του φάσματος σε 32 τμήματα (sub-bands)
- Εφαρμόζεται στο σήμα ένας ταχύς μετασχηματισμός Fourier (Fast Fourier Transform)
- Ένα ψύχο-ακουστικό μοντέλο εφαρμόζεται στο μετασχηματισμένο σήμα για να υπολογιστεί το ελάχιστο επίπεδο θορύβου που γίνεται αντιληπτό από το μέσο ακροατή

Το MPEG-1 προβλέπει δύο ηχητικά κανάλια. Αυτά μπορούν να απλά (μονοφωνικά, διπλά (δύο μονοφωνικά κανάλια), απλά στερεοφωνικά (ένα κανάλι μεταφέρει το αριστερό ηχητικό σήμα και το άλλο το δεξιό) ή από κοινού στερεοφωνικά (joint stereo, το ένα κανάλι μεταφέρει το άθροισμα και το άλλο τη διαφορά των σημάτων). Το πρότυπο χρησιμοποιεί 16bits για την κωδικοποίηση των δειγμάτων ενώ η συχνότητα δειγματοληψίας είναι 44.1kHz, 48kHz ή 32kHz.

Οι επιδόσεις κάθε στρώματος είναι:

- MPEG-Audio Layer-1: επιτρέπει την κατασκευή πολύ απλών κωδικοποιητών και αποκωδικοποιητών θυσιάζοντας βέβαια ένα μέρος της ποιότητας. Ως αποτέλεσμα, η ηχητική ποιότητα είναι μέτρια ενώ το bandwidth που απαιτείται είναι αυξημένο: 192 ή 256Kbps ανά κανάλι.
- MPEG-Audio Layer-2: ο αλγόριθμος αυτή της κατηγορίας έχει βελτιστοποιηθεί για ένα εύρος ζώνης 96 ή 128Kbps ανά μονοφωνικό κανάλι. Η ποιότητα είναι εφάμιλλη του CD.
- MPEG-Audio Layer-3: έχει την καλύτερη επίδοση από τα τρία στρώματα. Η ποιότητα του είναι υποδεέστερη αλλά πολύ κοντά σε αυτή του CD.

Το βασικό του πλεονέκτημα είναι ότι απαιτεί μόνο 64Kbps. Σε σύγκριση με MPEG-Audio Layer-2 στα 64Kbps λειτουργεί πολύ καλύτερα.

Να σημειωθεί ότι αυτά τα στρώμα έχουν προς τα πίσω συμβατότητα, δηλαδή μια συσκευή MPEG-Audio Layer-3 αποκωδικοποιεί και δεδομένα κωδικοποιημένα κατά MPEG-Audio Layer-1 ή 2.

- **MPEG-2**

Το MPEG-2 βρίσκεται ακόμα σε εξέλιξη. Στόχος του είναι η μετάδοση Ψηφιακής τηλεόρασης και video telephony μέσω οπτικών ινών, δορυφορικών καναλιών, ISDN και άλλων δικτύων. Το bandwidth που απαιτεί βρίσκεται στα 3-10 Mbits/s. Βασική επέκταση που επιφέρει στο MPEG-2 είναι η εισαγωγή διακριτών surround-sound καναλιών καθώς και μερικών καναλιών χαμηλής ταχύτητας για σχολιασμό η υποστήριξη πολλών γλωσσών.

## 4. Video

### 4.1 Video και Εφαρμογές

Η παρουσίαση των πρώτων εφαρμογών για προσωπικούς υπολογιστές που έκαναν χρήση video δημιούργησε ένα πολύ μεγάλο ενδιαφέρον. Παρ' όλα αυτά, ο τρόπος με τον οποίο θα χρησιμοποιηθεί αυτή η νέα τεχνολογία στο μέλλον δεν είναι ακόμα ξεκάθαρος[2]. Οι υπάρχουσες εφαρμογές μπορούν να χωριστούν σε δυο κατηγορίες: αναπαραγωγή αποθηκευμένου οπτικοακουστικού υλικού και πραγματικού χρόνου οπτικοακουστική επικοινωνία.

Η πρώτη κατηγορία είναι πιο καλά καθορισμένη και ήδη ευρέως χρησιμοποιούμενη. Υπάρχουν πολλές εκπαιδευτικές και ψυχαγωγικές εφαρμογές στις οποίες μέρος της πληροφορίας βρίσκεται σε μορφή video που αναπαράγεται ανάλογα με τις ανάγκες της εφαρμογής. Τέτοιου είδους δικτυακές εφαρμογές είναι ακόμα περιορισμένες λόγω τεχνολογικών προβλημάτων. Μια πιθανή εφαρμογή είναι η χρήση εξυπηρετητών που θα αποθηκεύουν μεγάλες βιβλιοθήκες video-clips και θα τα μεταδίδουν κατόπιν αιτήσεως του χρήστη (video-on-demand).

Η οπτικοακουστική επικοινωνία μπορεί να είναι ένας-προς-έναν, όπως για παράδειγμα η συνομιλία δύο ατόμων μέσω υπολογιστή που είναι εφοδιασμένος με κάμερα και συνδεδεμένος σε δίκτυο. Σε σύγκριση με τις εξειδικευμένες συσκευές τηλεδιάσκεψης, αυτά τα συστήματα προσωπικών υπολογιστών υστερούν σημαντικά σε ποιότητα. Η ανάλυση της εικόνας, το βάθος χρώματος και ο ρυθμός ανανέωσης των πλαισίων είναι πολύ μικρά. Κατά συνέπεια, για εφαρμογές που η οπτική επαφή είναι σημαντική, για παράδειγμα η συνέντευξη ενός νέου υπαλλήλου, τα συστήματα αυτά δεν είναι κατάλληλα. Υπάρχουν όμως περιπτώσεις που η ποιότητα βρίσκεται σε δεύτερη μοίρα και τα συστήματα αυτά βρίσκουν εφαρμογή. Η ενημέρωση ενός διευθυντή από υπαλλήλους που εργάζονται σε διαφορετικά σημεία μέσω ενός κοινού χώρου εργασίας (shared workspace) είναι ένα παράδειγμα τέτοιας εφαρμογής.

Σε πολλές περιπτώσεις το μοντέλο επικοινωνίας ένας-προς-πολλούς είναι πιο κατάλληλο. Για παράδειγμα, οι υπάλληλοι μιας εταιρείας ή οι φοιτητές ενός πανεπιστημίου μπορούν να παρακολουθούν ένα σεμινάριο που γίνεται σε κάποια άλλη τοποθεσία λαμβάνοντας την εικόνα μέσω δικτύου και αναπαράγοντας της στον υπολογιστή τους. Αυτές οι εφαρμογές έχουν το μεγάλο μειονέκτημα ότι για να σταλεί το σήμα σε πολλούς παραλήπτες απαιτούνται ταχύτατα δίκτυα και εξυπηρετητές.

## 4.2 Σύλληψη Video

Μια οθόνη τηλεόρασης μπορεί να μοιάζει φαινομενικά με αυτής του υπολογιστή αλλά στην ουσία υπάρχουν πολλά διαφορετικά στοιχεία. Τα υποσυστήματα γραφικών των υπολογιστών συνθέτουν την εικόνα με την τριάδα βασικών χρωμάτων RGB[3]. Αντίθετα, στις τηλεοράσεις το σύνθετο σήμα αποτελείται από τη φωτεινότητα και τις δυο χρωματικές συνιστώσες. Επιπλέον, η σάρωση στην τηλεόραση είναι πλεκτή ενώ οι ρυθμοί ανανέωσης και το μέγεθος του πλαισίου ποικίλουν από πρότυπο αλλά σε κάθε περίπτωση διαφέρουν από αυτά του υπολογιστή. Άρα για να γίνει η μετατροπή του τηλεοπτικού σήματος ή του σήματος ή του σήματος video σε μορφή κατάλληλη για υπολογιστή απαιτείται ειδικό υλικό.

Το υλικό αυτό πέρασε διάφορα στάδια εξέλιξης κάθε ένα εκ των οποίων κυριάρχησε έναντι των προκατόχων του. Επίσης, είναι φανερό η μετακίνηση προς τα διεθνή πρότυπα.

## 4.3 Τεχνικές Συμπίεσης

Σε μία εικόνα εντοπίζεται και απορρίπτεται το χωρικό πλεόνασμα πληροφορίας. Στην κινούμενη εικόνα, υπάρχει ένα ακόμα είδος πλεονάσματος το χρονικό πλεόνασμα[1]. Πιο συγκεκριμένα, όταν κινείται ένα αντικείμενο τα διαδοχικά πλαίσια μοιάζουν σημαντικά. Κάποια τμήματα των πλαισίων δεν επηρεάζονται καθόλου από την κίνηση ενώ κάποια άλλα πιθανόν να αλλάζουν απλώς θέση με μικρή ή και καμία αλλαγή του περιεχομένου τους.

Ένας αλγόριθμος συμπίεσης κινούμενης εικόνας μπορεί να στηρίζεται μόνο στην εξάλειψη του χωρικού πλεονάσματος ή να συνδυάζει εξάλειψη και των δύο ειδών πλεονασμάτων. Το MJPEG είναι ένα παράδειγμα της πρώτης κατηγορίας και το MPEG της δεύτερης. Πλεονέκτημα των αλγορίθμων της πρώτης κατηγορίας είναι η ευκολία επέμβασης στην εικόνα σε επίπεδο πλαισίου και η ανθεκτικότητα σε λάθη κατά την μετάδοση μέσω δικτύου. Αυτό συμβαίνει γιατί κάθε πλαίσιο είναι ανεξάρτητο από τα υπόλοιπα και κάθε λάθος επηρεάζει μόνο αυτό. Αντίθετα σε τεχνικές όπως το MPEG υπάρχει συσχέτιση κάθε πλαισίου με τα προηγούμενα του, οπότε και είναι δύσκολο να γίνει η εξαγωγή των πλαισίων και κάθε λάθος έχει επιπτώσεις σε όλα τα συσχετιζόμενα πλαίσια. Αναμφισβήτητα, η εξάλειψη και του χωρικού και του χρονικού πλεονάσματος οδηγεί σε σημαντικά μεγαλύτερους λόγους συμπίεσης γι' αυτό και το ενδιαφέρον εστιάζεται σε αυτούς.

#### 4.4 Το πρότυπο συμπίεσης MPEG

Τα αρχικά **MPEG** προέρχονται από τις λέξεις **Moving Picture Experts Group** (Ομάδα Ειδικών στην Κινούμενη Εικόνα)[5]. Πρόκειται για μία επιτροπή που δρα στα πλαίσια του Διεθνούς Οργανισμού τυποποίησης. Επίσημα είναι γνωστή σαν ISO/IEC JTC1/SC29/WG11. Ιδρύθηκε το 1988 και είναι μέλος του JTC1 (Joint ISO/IC Technical Committee on Information Technology - Ενωμένη Τεχνική επιτροπή στην Τεχνολογία της Πληροφορικής). Ο συντονιστής της επιτροπής MPEG είναι ο Leonardo Chiariglione γνωστός σαν ο «πατέρας» του MPEG. Η επιτροπή πραγματοποιεί 3 -4 συναντήσεις το χρόνο σε διάφορες πόλεις του κόσμου όπου συζητούνται οι εξελίξεις της ερευνητικής δουλειάς που έχει γίνει ενδιάμεσα, θέτονται στόχοι και προθεσμίες και διατυπώνονται οι προδιαγραφές πάνω στις οποίες οι εταιρίες θα αναπτύξουν τα προϊόντα.

Το όνομα MPEG έχει επικρατήσει όμως να αναφέρεται και στην οικογένεια των τυποποιήσεων (standards) που δημιουργήθηκαν από την ομάδα MPEG και χρησιμοποιούνται για τη μετάδοση οπτικών και ηχητικών δεδομένων σε ψηφιακή συμπίεσμένη μορφή. Η οικογένεια MPEG περιλαμβάνει τα standards **MPEG-1, MPEG-2, MPEG-3, MPEG-4** και το επερχόμενο **MPEG-7**.

Πιο αναλυτικά:

- **MPEG-1 Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s**

Αναπτύχθηκε για την αποθήκευση και ανάκτηση κινούμενης εικόνας και ήχου σε ψηφιακά μέσα με ρυθμό μετάδοσης μέχρι 1,5 Mbits/sec. Η εικόνα έχει ανάλυση 352x240 pixels (NTSC) ή 352x288 pixels (PAL) και η ποιότητά της είναι σε επίπεδα VHS video. Χρησιμοποιείται κυρίως για την αποθήκευση video σε CD-ROM, Video-CD και CD-i και όπου αλλού χρειάζεται μικρό (σε σχέση με το MPEG-2) bandwidth. Το MPEG-1 μπορεί να χρησιμοποιηθεί σε εφαρμογές με ρυθμό μετάδοσης 4-5 Mbits/sec, αλλά τα αποτελέσματα δεν είναι τόσο καλά, όσο στην κανονική περιοχή λειτουργίας του.

- **MPEG-2 Generic coding of moving pictures and associated audio information**

Αναπτύχθηκε για εφαρμογή στην ψηφιακή τηλεόραση.

Η βασική ανάλυση της εικόνας ακολουθεί το τηλεοπτικό πρότυπο CCIR-601 (broadcast quality - ποιότητα εκπομπής) δηλαδή 704x480 pixels (NTSC) ή 704x576 pixels (PAL) και υποστηρίζει εικόνα πλεκτής σάρωσης (interlaced).

Ο ρυθμός μετάδοσης κυμαίνεται από 3 ως 10 Mbits/sec. Οι εφαρμογές του είναι στην καλωδιακή τηλεόραση (Cable TV), στη δορυφορική (Direct Broadcasting Satellite TV) αλλά αναμένεται να επεκταθεί και στην επίγεια τηλεόραση. Επίσης χρησιμοποιείται στην αποθήκευση κινηματογραφικών ταινιών στα DVD (Digital Video Disk).

- **MPEG-3**

Είναι προσανατολισμένο στην τεχνολογία της **Τηλεόρασης Υψηλής Ευκρίνειας** (HDTV - High Definition TV) αλλά εγκαταλείφθηκε αφού διαπιστώθηκε ότι το MPEG-2 μπορεί με κάποιες αλλαγές στη σύνταξη των προδιαγραφών να χρησιμοποιηθεί το ίδιο καλά στη HDTV. Έτσι η δουλειά που είχε γίνει πάνω στο MPEG-3 ενσωματώθηκε στο MPEG-2.

- **MPEG-4 Coding of audio-visual objects**

Ο όρος audio visual objects (AV-objects) είναι γενικός και σημαίνει διάφορες οντότητες που απαρτίζουν την εικόνα και οι οποίες μπορούν κωδικοποιητής και αποκωδικοποιητής να χειρισθούν αυτόνομα και ανεξάρτητα από τις υπόλοιπες. Με τον όρο οντότητες πάλι εννοούμε σχήματα και ήχους, φυτικούς ή computer generated που χρησιμοποιούνται για να αναπαραστήσουν άλλα ομοειδή αντικείμενα. Είναι ένα standard για εφαρμογές επικοινωνίας πολυμέσων ( multimedia communications) δηλαδή εφαρμογές όπως video-phone, video-conference, video e-mail, electronic news και πολλές άλλες. Η ανάλυση της εικόνας είναι 176x144 pixels σε σχετικά χαμηλούς ρυθμούς μετάδοσης που κυμαίνονται ανάμεσα στα 4.8 και 64 Kbits/sec, κατάλληλα δηλαδή για μετάδοση σε δίκτυα με μικρό διαθέσιμο bandwidth ανά συνδρομητή, όπως το Internet.

- **MPEG-7**

Είναι ένα standard κωδικοποίησης με αναπαράσταση περιεχομένου (content representation), για την αναζήτηση πληροφοριών σε εφαρμογές πολυμέσων.

Να σημειωθεί ότι οι αναλύσεις (resolution) της εικόνας που αναφέρθηκαν παραπάνω δεν είναι περιοριστικές αλλά αναφέρονται στους περιορισμούς που έχουν τεθεί για να κρατηθούν σε λογικά επίπεδα η πολυπλοκότητα των κωδικοποιητών και αποκωδικοποιητών και ο όγκος δεδομένων. Ο περιορισμός αυτός ονομάζεται **CPB** (Constrained Parameters Bitstream) και ορίζει τις διαστάσεις που πρέπει να έχουν τα MPEG σήματα, κάτι σαν ένα standard format.

Παρόλα αυτά μπορεί να γίνει κωδικοποίηση και σε υψηλότερες αναλύσεις απλώς δεν υπάρχει εγγύηση ότι θα μπορούν να χρησιμοποιηθούν με όλους τους διαθέσιμους αποκωδικοποιητές, άσχετα αν ακολουθούν τους κανόνες του MPEG. Έτσι το MPEG-2 π.χ. μπορεί να φτάσει ανάλυση (resolution) 1920x1080 και το MPEG-1 4095x4095. Επίσης το γεγονός ότι τα σήματα MPEG εμφανίζονται σε δύο διαφορετικές αναλύσεις (διαστάσεις) εικόνας οφείλεται στην ύπαρξη δύο συστημάτων για το αναλογικό σήμα, τα PAL και NTSC, με δειγματοληψία των οποίων προκύπτουν τα σήματα MPEG. Ακόμα δεν έχει γίνει σημαντική πρόοδος στην κατεύθυνση της δημιουργίας πηγών (κάμερες κτλ.) που θα παράγουν σήμα MPEG απευθείας.

Τα δύο τελευταία στάδια του MPEG είναι ακόμα υπό ανάπτυξη και δεν έχουν γίνει ακόμα standards (το MPEG-4 δεν έχει γίνει ακόμα standard και το MPEG-7 είναι ακόμα στα προκαταρκτικά σχέδια) ενώ τα MPEG-1 και MPEG-2 έχουν τεθεί ήδη σε εκτεταμένη εφαρμογή. Σε γενικές γραμμές το MPEG-1 έχει αντικατασταθεί από το MPEG-2 το οποίο είναι μια βελτίωση του, που προσφέρει καλύτερη εικόνα και λόγους συμπίεσης και είναι συμβατό με το MPEG-1 (**backwards compatible**). Αυτό σημαίνει ότι συσκευές MPEG-2 μπορούν να χειριστούν σήματα συμπιεσμένα με MPEG-1 (αλλά όχι το αντίστροφο).

#### 4.5 Είδη πλαισίων (frames)

Δύο βασικές προσεγγίσεις στην κωδικοποίηση της κινούμενης εικόνας είναι οι παρακάτω. Η πρώτη είναι βασισμένη στη μέθοδο συμπίεσης ακίνητης εικόνας JPEG και η δεύτερη στην τμηματική πρόβλεψη κίνησης (block motion compensation)[5]:

- **Ενδο-πλαισιακή Κωδικοποίηση (Intra-frame Coding):**

Σύμφωνα με αυτή την προσέγγιση η κάθε εικόνα (frame) αντιμετωπίζεται σαν αυτόνομη μονάδα και κωδικοποιείται ανεξάρτητα από τις υπόλοιπες, οπότε το τελικό σήμα είναι μία σειρά από διακριτές ακίνητες εικόνες.

- **Δια-πλαισιακή Κωδικοποίηση (Inter-frame Coding):**

Λαμβάνονται κατά την κωδικοποίηση υπ' όψη οι πιθανές ομοιότητες μεταξύ των πλαισίων και κωδικοποιείται η διαφορά τους με χρήση του block motion compensation. Έτσι στο τελικό σήμα υπάρχει μία εξάρτηση μεταξύ των πλαισίων αφού για να αποκωδικοποιηθεί κάποιο πιθανώς να πρέπει να ληφθούν πληροφορίες και από κάποιο άλλο (προηγούμενο ή επόμενο). Γίνεται έτσι πολύ μεγαλύτερη συμπίεση, αφού μόνο οι διαφορές μεταξύ των πλαισίων κωδικοποιούνται.

Σε ένα σήμα (bit-stream) MPEG υπάρχουν τριών ειδικών πλαίσια:

- **I (Intra frames):**

Το είδος αυτό των πλαισίων κάνει χρήση του intra frame-coding. Τα πλαίσια **τύπου I** είναι τα μόνα που είναι κωδικοποιημένα στο σύνολό τους και η αποκωδικοποίηση μπορεί να γίνει χωρίς αναφορά σε κάποιο άλλο. Είναι κατά συνέπεια τα μεγαλύτερα σε μήκος και αποτελούν σημεία αναφοράς κατά την τυχαία προσπέλαση ενός σήματος. Επειδή η παρουσία τους είναι απαραίτητη σε σημείο χρονικής αναφοράς και για να αποφευχθεί η διάδοση των σφαλμάτων που δημιουργούν τα P πλαίσια επιβάλλεται να μεταδίδονται ανά τακτά χρονικά πλαίσια. Έτσι υπάρχει ένα I πλαίσιο τουλάχιστον κάθε 15 πλαίσια (δηλαδή δύο φορές το δευτερόλεπτο αν η συχνότητα είναι 30 Hz).

Η διαδικασία της κωδικοποίησης ενός I πλαισίου φαίνεται στον παρακάτω σχήμα. Η εικόνα χωρίζεται σε macroblocks και για κάθε block ξεχωριστά εφαρμόζεται DCT, Κβαντοποίηση, Zig-Zag Scanning, Run-Length-Encoding και Huffman Encoding (οι μέθοδοι αυτοί θα εξεταστούν ξεχωριστά παρακάτω).

- **P (Predicted frames):**

Τα πλαίσια **τύπου P** είναι βασισμένα σε ένα προηγούμενο I ή P πλαίσιο. Με τη βοήθεια του motion compensation προβλέπουν τη νέα θέση όποιων macroblocks έχουν απλά μετακινηθεί και κωδικοποιούν τον αριθμό του macroblock και ένα διάνυσμα κίνησης. Με τη σειρά τους μπορούν να αποτελέσουν και αυτά σημείο αναφοράς για επόμενα πλαίσια και αυτός είναι και ο λόγος που συμβάλλουν στην εισαγωγή και διάδοση σφαλμάτων, αφού η διαδικασία της πρόβλεψης κίνησης δεν μπορεί να είναι 100% ακριβής. Δεν έχουν το μέγεθος των I πλαισίων γιατί δεν έχουν περιγράψει με την ίδια ακρίβεια, δηλαδή παρουσιάζουν μεγαλύτερο ποσοστό συμπίεσης. Η διαδικασία λοιπόν της κωδικοποίησης τους, όπως φαίνεται και από το παρακάτω σχήμα είναι: σύγκριση macroblocks και δημιουργία ενός γραμμικού συνδυασμού αυτών που παρουσιάζουν σημαντική ομοιότητα, δημιουργία motion vector, μετασχηματισμός DCT σε κάθε block του νέου macroblock, Κβαντοποίηση, Run-Length-Encoding και το αποτέλεσμα κωδικοποιείται με κωδικοποίηση Huffman, όπως και στα I πλαίσια.

- **B (Bi-directional frames):**

Τα πλαίσια **τύπου B** είναι πλαίσια που δημιουργούνται λαμβάνοντας το μέσο όρο σε επίπεδο macroblock ενός προηγούμενου και ενός επόμενου πλαισίου I και P (ένα από το κάθε είδος).



Δε συντελούν τόσο πολύ στη διάδοση των σφαλμάτων γιατί δεν χρησιμοποιούνται ως σημεία αναφοράς και επιπλέον μειώνουν σημαντικά το σφάλμα παίρνοντας το μέσο όρο από δύο πλαίσια. Μπορούμε να πούμε ότι ο 'κύκλος της ζωής' τους περιορίζεται μόνο σε αυτά και δεν επεκτείνεται με το να κληροδοτούν πληροφορίες σε άλλα πλαίσια, κάτι που πολλές φορές σε συνδυασμό και με την υπολογιστική πολυπλοκότητα που απαιτούν για την κωδικοποίηση και αποκωδικοποίηση τα κάνει με επιθυμητά από τους κατασκευαστές. Η διαδικασία της κωδικοποίησης περιλαμβάνει συνδυασμό των αντίστοιχων macroblocks που παρουσιάζουν μικρές διαφορές με τα αντίστοιχα των πλαισίων αναφοράς (προηγούμενο και επόμενο) δηλαδή αφαίρεση του μέσου όρου των άλλων δύο από το τρέχον πλαίσιο, συνδυασμό των διανυσμάτων κίνησης των πλαισίων αναφοράς (που συνδυάζονται όπως και τα αντίστοιχα macroblocks, δηλαδή λαμβάνεται ο μέσος όρος τους) και στη συνέχεια την ίδια διαδικασία με τα I και P πλαίσια για την κωδικοποίηση του macroblock που προκύπτει.

Ο κύριος λόγος ύπαρξης των B- πλαισίων είναι η κάλυψη της περίπτωσης κάποιες πληροφορίες της εικόνας να υπάρχουν σε επόμενα πλαίσια και να μην υπάρχουν στα προηγούμενα. Συνεπώς η πρόβλεψή τους με τα P πλαίσια θα ήταν αδύνατη. Σαν παράδειγμα μπορούμε να αναφέρουμε ία πόρτα που ανοίγει ξαφνικά. Η πληροφορία για το τι βρίσκεται πίσω από την πόρτα υπάρχει στα επόμενα πλαίσια και όχι στα προηγούμενα και για να εμφανιστεί και στο τρέχον πλαίσιο πρέπει να ληφθούν σαν σημεία αναφοράς και το προηγούμενο και το επόμενο.

Αφού τα πλαίσια P κατασκευάζονται με βάση τα I και τα B με βάση τα I και P είναι προφανές ότι τα I πρέπει να έχουν σταλεί πριν τα αντίστοιχα P. Επίσης και τα P αλλά και τα I πρέπει να έχουν σταλεί πριν από αντίστοιχα B, παρόλο που στη μετάδοση αυτά παρεμβάλλονται ανάμεσα τους. Η σειρά με την οποία απεικονίζονται και η σειρά με την οποία αποστέλλονται τα πλαίσια (που προφανώς δεν είναι η ίδια) φαίνεται στο παρακάτω σχήμα:

Η συνηθέστερη διάταξη των πλαισίων σε ένα σήμα MPEG είναι η παρακάτω[2]. Πολλές φορές παρεμβάλλονται περισσότερα B πλαίσια και τα I πλαίσια απέχουν περισσότερο μεταξύ τους (αλλά αυτό υποβαθμίζει την ποιότητα της εικόνας γιατί τα σφάλματα διαδίδονται περισσότερο).

Η μικρότερη μονάδα που μπορεί να αποκωδικοποιηθεί ανεξάρτητα ονομάζεται GOP (Group of Pictures) και περιέχει όλα τα I,P,B, πλαίσια που χρειάζονται για την αποκωδικοποίηση, χωρίς να γίνονται αναφορές σε άλλο GOP.

Στο παρακάτω πίνακα φαίνεται η αναλογία πλαισίων I,P,B σε αν σήμα MPEG:

Είδος εικόνας	Bit-rate	I	P	B	Μέσος όρος
MPEG-1	(1.15Mbit/s)	150,000	50,000	20,000	38,000
MPEG-2	(4.00Mbit/s)	400,000	200,000	80,000	130,000

Σχήμα 3: Αναλογία πλαισίων I,P,B σε ένα σήμα MPEG

#### 4.6 Ψηφιακό interactive βίντεο (Digital video interactive DVI)

Το DVI της Intel είναι ένα ντε φάκτο πρότυπο συμπίεσης[1]. Το CD-ROM μπορεί να αποθηκεύσει μόνο μια ώρα ακουστικού CD, λιγότερες από 500 υψηλής ανάλυσης στατικές εικόνες ή 30 δευτερόλεπτα ασυμπίεστου, πλήρους κίνηση ψηφιακού βίντεο. Παρόλα αυτά τώρα υπάρχουν χιλιάδες εμπορικές διαθέσιμες εφαρμογές χρησιμοποιώντας CD-ROM, ένας μεγάλος αριθμός αφορά κείμενο μόνο ή τις πιο πολλές φορές γραφικά. Το DVI είναι μια εμπορική επιτυχής τεχνολογία η οποία έχει φέρει το βίντεο στο PC. Για παράδειγμα ένα DVI CD-ROM μπορεί να αποθηκεύσει 20 λεπτά βίντεο ή 1000 από 1 λεπτό ήχου και ακόμη θα υπάρχει χώρος για 50,000 σελίδες κειμένου.

Το DVI είναι ένας ασύμμετρος αλγόριθμος συμπίεσης: ειδικευμένα συστήματα συμπιέζουν το κινούμενο βίντεο με έναν τρόπο μη ελεγχόμενο από τον υπολογιστή, ενώ χαμηλού κόστους επιτραπέζια συστήματα παίζουν τις εικόνες σε πραγματικό χρόνο. Το DVI βίντεο βασίζεται σε δύο αλγόριθμους: παραγωγή κατά επίπεδο βίντεο (production-level-video-PLV) και το πραγματικού χρόνου βίντεο (real time video RTV). Το PLV παρέχει υψηλότερη ποιότητα αλλά δεν δουλεύει σε πραγματικό χρόνο. Το RTV αποδίδει καλά σε μέγεθος ίσο με το ένα τέταρτο της κανονικής οθόνης αλλά δεν αποδίδει για πλήρη οθόνη. Για παράδειγμα το ActionMedia II της IBM μπορεί να παίζει PLV αλλά μπορεί μόνο να κάνει εγγραφή RTV. Τα σύγχρονα έχουν ρυθμό μετάδοσης 150 KBps, γι' αυτό το λόγο είναι κατάλληλα για την αποθήκευση DVI-κωδικοποιημένου υλικού. Επίσης είναι κατάλληλο για δικτυακές εφαρμογές πολυμέσων.

Οι εφαρμογές μπορούν να αποθηκευτούν σε compact disk ή σε σκληρό δίσκο. Το CD-ROM τυπικά περιλαμβάνει βιβλιοθήκες DVI συστήματος και όλα τα βίντεο κλιπς τα οποία απαιτούνται από τις εφαρμογές. Η τεχνολογία DVI έχει κιόλας περάσει στην 3ή γενιά προϊόντων. Τα αρχικά συστήματα DVI αποτελούνταν από ένα PC το οποίο περιείχε 3 custom board και συνοδευτικό λογισμικό. Σε λίγα χρόνια, (Μπορεί σε 2) το DVI θα διατίθεται πάνω στη μητρική του PC. Η αρχική πλατφόρμα DVI (γνωστή ως DVI - end-user system) ήταν ένα IBM PC/AT ή μια συμβατή μηχανή, με αναλογική RGB οθόνη, ενισχυτή ήχου και ηχεία, έναν οδηγό CD-ROM και 3 DVI board.

Το πρώτο board είναι ένα βίντεο board το οποίο περιέχει δύο επεξεργαστές βίντεο. Τα άλλα boards υποστηρίζουν ήχο και CD-ROM interface, αντιστοίχως. Ο επεξεργαστής pixel στην κάρτα βίντεο δημιουργεί τις μορφές bitmap και ο επεξεργαστής οθόνης στην ίδια κάρτα βίντεο τις αναπαριστά στην RGB οθόνη. Ο επεξεργαστής pixel τρέχει στα 12,5 MIPS ή και ψηλότερα και προσκομίζει μικροκωδικοποιημένους αλγορίθμους από τη μνήμη RAM του 1 MB. Ο επεξεργαστής pixel υποστηρίζει δημιουργία και επεξεργασία ψηφιακού βίντεο και γραφικών και περιέχει ειδικού σκοπού υλικό για να βοηθήσει στην αποκωδικοποίηση σε πραγματικό χρόνο των βίντεο που έχουν υποστεί υψηλή συμπίεση. Ο επεξεργαστής οθόνης έχει τη δυνατότητα εμφάνισης bitmaps σε διάφορες αναλύσεις: 256x240 ανάλυση για πλήρη οθόνη για βίντεο και μέχρι 1024x512 για στατικές εικόνες ή για βίντεο που έχουν μέγεθος μικρότερο της πλήρους οθόνης. Ο επεξεργαστής οθόνης υποστηρίζει 8-bit, 16-bit και 32-bit pixel depth.

#### 4.7 Πρότυπα ITU-T H.200

Το 1984 το CCITT Study Group XV καθιέρωσε μια εξειδικευμένη ομάδα στην κωδικοποίηση για οπτική τηλεφωνία για να αναπτύξει συστάσεις για μετάδοση video στα  $m \times 384$  Kbps ( $m=1-5$ ). Αργότερα στην περίοδο μελέτης, τυποποίηση στα  $p \times 64$  Kbps ( $p=1-30$ ) επίσης μελετήθηκε, καθώς η τεχνολογία συμπίεσης βελτιώθηκε και έτσι έκανε εφικτή τη λειτουργία του video σε χαμηλότερους ρυθμούς[8]. Οι συστάσεις που εμφανίστηκαν είναι γνωστές σαν «υποδομή για οπτικοακουστικές υπηρεσίες». ITU-T/CCITT σύσταση H261, «κωδικοποίηση video για οπτικοακουστικές υπηρεσίες στα  $p \times 64$  Kbps», καθορίζει έναν κώδικα για συμπιεσμένο ψηφιακό video ο οποίος ολοκληρώθηκε και εγκρίθηκε το Δεκέμβριο του 1990. Μία από τις εφαρμογές αυτού του τύπου για video-τηλέφωνο και video-διάσκεψη. Συνεπώς, ο αλγόριθμος κωδικοποίησης video πρέπει να είναι ικανός να λειτουργεί σε πραγματικό χρόνο με τη μικρότερη καθυστέρηση. Για  $p=1$  ή 2, μόνο επιτραπέζια άμεση οπτική επικοινωνία («videophony») μπορεί να επιτευχθεί πρακτικά. Πιο πολύπλοκες εικόνες, για παράδειγμα, στην περίπτωση της video-διάσκεψης απαιτούν  $p > 6$ .

Εκτός από τη σύσταση πλαισίου εργασίας H.200 και τη σύσταση H.261, το ITU-TS έχει υιοθετήσει δυο τύπους για υποστήριξη της video-τηλεφωνίας και video-διάσκεψης.

- Το H.221 καθορίζει μια κατασκευή πλαισίων για την υποστήριξη οπτικοακουστικών τηλε-υπηρεσιών σε 64Kbps κανάλια
- Το H.222 καθορίζει μια κατασκευή πλαισίων για την υποστήριξη οπτικοακουστικών τηλε-υπηρεσιών σε περιβάλλον ISDN στα  $p \times 64$  Kbps  $p=1,2,\dots,30$  (για επιλεγμένες τιμές).

Σαν ομάδα το H.221 και το H.222 καθορίζουν μια κατασκευή πλαισίων για τη υποστήριξη οπτικοακουστικών τηλε-υπηρεσιών σε περιβάλλον ISDN χρησιμοποιώντας είτε i) ένα μονό ή πολλαπλό B ή H0 κανάλια ή ii) ένα μονό H11 ή H12 κανάλι. Η δυνατότητα κατανομής ρυθμού bit παρέχει στα κανάλια λειτουργία σε συνολική ταχύτητα  $px64$  Kbps για  $p=1,2,3,4,6,12,18,23,24,30$ . Επιπλέον υπάρχει η ικανότητα για συγχρονισμό πολλαπλών 64 Kbps ή 384 Kbps καναλιών: για παράδειγμα δυο 64 Kbps κανάλια μπορούν να δημιουργηθούν πάνω σε διαφορετικές διαδρομές στο δίκτυο και μπορούν να ανασυγχρονιστούν στο λαμβανόμενο τερματικό.

Στο περισσότερο υλικό video υπάρχει μόνο περιορισμένη αλλαγή από το ένα πλαίσιο στο άλλο. Το H.261 συγκρίνει κάθε πλαίσιο με το προηγούμενο για να εκτελέσει μια προφητική ανάλυση συμπίεσης κίνησης. Μερικά υψηλής τεχνολογίας video-chips είναι προγραμματισμένα και υποστηρίζουν JPEG, MPEG-1, και ITU-T/CCITT H.261 video συμπίεση και αποσυμπίεστη. Μερικοί αμφισβητούν την αξία του να περιλαμβάνεται ένας τύπος συμπίεσης σαν τμήμα ενός τύπου μετάδοσης ενώ άλλοι αναρωτιούνται αν είναι προτιμητέο να ληφθεί μια ευέλικτη προσέγγιση συμπίεσης που επιτρέπει στο χρήστη να επιλέγει επίπεδα συμπίεσης σύμφωνα με τις απαιτήσεις της εφαρμογής. Ας δούμε πιο αναλυτικά τα πρότυπα αυτής της οικογένειας.

- **H.200:**

Ο σκοπός της σύστασης H.200 «πλαίσιο εργασίας για συστάσεις για οπτικοακουστικές υπηρεσίες» είναι να καθορίσει ένα σύνολο οπτικοακουστικών υπηρεσιών το οποίο έχει σύμφωνες απαιτήσεις και χαρακτηριστικά. Ο πίνακας 3 που ακολουθεί παρέχει το σύνολο των υπηρεσιών που περιλαμβάνονται στις προσπάθειες τυποποίησης ITU-TS.

- **H.221:**

Το H.221 είναι ο πιο σημαντικός τύπος ελέγχου. Καθορίζει τη κατασκευή πλαισίων για οπτικοακουστικές υπηρεσίες σε ένα ή πολλαπλά B ή H0 κανάλια ή μονά H11 ή H12 κανάλια σε ρυθμούς μεταξύ 64 και 1920 Kbits/sec. Επιτρέπει το συγχρονισμό πολλαπλών 64 ή 384Kbits/sec συνδέσεων και δυναμικό έλεγχο στην υποδιαίρεση του καναλιού μετάδοσης των 64 μέχρι 1920 Kbps σε μικρότερα υποκανάλια κατάλληλα για φωνή, video, δεδομένα και σήματα ελέγχου. Είναι κυρίως σχεδιασμένο για χρήση με συγχρονισμένες συνδέσεις πολυμέσων όπως video-διάσκεψη. Επίσης προσφέρει αρκετά πλεονεκτήματα:

- Λαμβάνει υπόψη τη σύσταση G.704 και X.30/I.361 ανάμεσα σε άλλες και επιτρέπει τη χρήση του υπάρχοντος hardware και software.

- Είναι απλό οικονομικό και ευέλικτο και έτσι μπορεί να εφαρμοστεί σε
- έναν απλό μικροεπεξεργαστή.
- Είναι μια σύγχρονη διαδικασία: ο ακριβής χρόνος μιας αλλαγής διαμόρφωσης είναι ο ίδιος για τον πομπό και τον δέκτη. Οι διαμορφώσεις μπορούν να αλλάξουν σε χρόνο 20ms.
- Δεν απαιτεί σύνδεσμο επιστροφής αφού η διαμόρφωση στέλνεται από μια επαναλαμβανόμενα μεταδιδόμενη λέξη κώδικα.
- Είναι ασφαλής στην περίπτωση μετάδοσης σφαλμάτων επειδή το σήμα διανομής του ρυθμού bit προστατεύεται από ένα διπλό διορθωτικό κώδικα.
- Επιτρέπει τον έλεγχο μιας υψηλότερης πολυπλεκτικής διαμόρφωσης, μέσα στην οποία εισχωρεί το βασικό 64 Kbit/sec κανάλι.
- Παρέχει μια ποικιλία ρυθμών bit στο χρήστη.

Αρχικό σύνολο	Μελλοντικό σύνολο
Στενής ζώνης video-τηλέφωνο	Video ταχυδρομείο
Ευρείας ζώνης video-τηλέφωνο	Videotext (με εικόνες και ήχους)
Στενής ζώνης video-τηλέφωνο	Video ανάκτηση
Ευρείας ζώνης video-διάσκεψη	Υψηλής ανάλυσης ανάκτησης εικόνας
Ηχογραφική τηλεδιάσκεψη	Υπηρεσίες διανομής

**Σχήμα 4:** Οπτικοακουστικές υπηρεσίες της ITU τυποποίησης:

- **H.222:**

Το H.222 «κατασκευή πλαισίων για 384-1920-Kbps κανάλια σε οπτικοακουστικές τηλε-υπηρεσίες» παρέχει ένα μηχανισμό σε πολύπλοκα σήματα όπως ήχος, video, δεδομένα και έλεγχο για οπτικοακουστικές ευρέως διαδεδομένες τηλε-υπηρεσίες χρησιμοποιώντας ένα mx384 Kbps κανάλι (m=1-5. Στο H.222 η σήμανση επιτυγχάνεται σε μια εσωτερικής ζώνης διαμόρφωση (in-band fashion) χρησιμοποιώντας έναν αριθμό bits στο αρχικό 64 Kbps κανάλι, γνωστό και σαν το αρχικό κανάλι ή το I-κανάλι. Εκτός από τη δημιουργία ευθυγραμμισμένων πλαισίων, αυτά τα bits επιτρέπουν κατανομή ρυθμού bit, απόκρυψη, και end-to-end παρακολούθηση ποιότητας. Άλλα 64 Kbps κανάλια σχετιζόμενα με ένα δοθέν (-κανάλι μπορεί να περιλαμβάνει ή όχι δηλωτικά bit, αναλόγως τη λειτουργία. Για παράδειγμα μια HO κλήση στα 384 Kbps θα απαιτεί μόνο ένα κανάλι σήμανσης, αφού τα εναπομείναντα 64 Kbps κανάλια ήδη συνδέονται με το I-κανάλι. Όταν δυο 64 Kbps διέρχονται από B-κανάλια συνδέονται το ένα με το άλλο για την παροχή ενός 128 Kbps καναλιού για μια video-διάσκεψη, η σήμανση είναι απαραίτητη να εκτελείται και στα δυο κανάλια.

- **H.230:**

Αυτός ο τύπος κυρίως ασχολείται με τα ελεγκτικά και τα δηλωτικά σήματα που χρειάζονται για την μετάδοση μιας σύγχρονου-πλαίσου ή μιας απαιτούμενης αστραπιαίας απόκρισης. Τέσσερις κατηγορίες ελεγκτικών και δηλωτικών σημάτων έχουν καθοριστεί, η πρώτη αφορά το video, η δεύτερη αφορά τον ήχο, η Τρίτη αφορά σκοπούς συντήρησης και η τελευταία αφορά τον απλό έλεγχο διασκέψεων (σήματα μεταδίδονται μεταξύ τερματικών και MCU's). Το H.230 είναι ένα πρωτόκολλο για έλεγχο και ανάμιξη video και ήχου και σκοπεύει σε απλές επεκτάσεις για point to point και ISDN video-τηλεφωνία.

- **H.242:**

Εξαιτίας του αυξανόμενου αριθμού εφαρμογών που χρησιμοποιούν στενής (3KHz) και ευρείας ζώνης (7KHz) ομιλία μαζί με video και δεδομένα σε διαφορετικούς ρυθμούς, ένα σχέδιο συνιστάται από την τυποποίηση για να επιτρέψει σε ένα κανάλι να εξυπηρετήσει ομιλία και προαιρετικά video και / ή δεδομένα σε μερικούς ρυθμούς και σε έναν αριθμό διαφορετικών καταστάσεων. Διαδικασίες σήμανσης για δημιουργία ενός συμβατού mode πάνω σε μια κλήση, για να αλλάζει θέση ανάμεσα στα modes και να επιτρέπει μια μεταβίβαση κλήσης, εξηγούνται σε αυτή την τυποποίηση.

Κάθε τερματικό θα μεταβιβάζει τις δυνατότητες του σε αλλά απομακρυσμένα τερματικά στο ξεκίνημα μιας κλήσης. Τα τερματικά έπειτα θα συνεχίσουν να δημιουργήσουν ένα κοινό mode λειτουργίας. Οι δυνατότητες ενός τερματικού είναι: δυνατότητες ήχου, δυνατότητες video, δυνατότητες ρυθμού μετάδοσης, δυνατότητες δεδομένων, δυνατότητες τερματικών σε περιορισμένα δίκτυα και δυνατότητες απόκρυψης και επέκτασης -BAS. Το H.242 έχει ένα πολυπλεκτικό πρωτόκολλο για να μεταφέρει πολλά διάφορα στενού εύρους σήματα ομιλίας και video. Έχει επίσης ένα πρωτόκολλο για διαπραγμάτευση των δυνατοτήτων ανάμεσα στα τερματικά.

- **H.231 και H.243:**

Δύο καθοριστικά νέα πρότυπα που βρίσκονται σε εξέλιξη απευθύνονται σε multipoint «γεφύρωση» δικτύων. Η multipoint διάσκεψη είναι η λογική προέκταση, ενθαρρύνοντας τη σχετικά απλή point to point διάσκεψη να χειριστεί διάφορα μέρη ταυτόχρονα. Το πρότυπο H.231 είναι σχεδιασμένο να ελέγχει τις συνδέσεις μεταξύ τριών ή περισσότερων ανόμοιων codecs με μονάδες multipoint ελέγχου (MCUs).

Το πρότυπο H.242 ελέγχει τις διαδικασίες ανάμεσα στα H.231 MCUs. Το κλειδί στο πρότυπο H.243 είναι οι κώδικες του bit rate σήματος διανομής που επιτρέπουν στους codecs να εναλλάσσονται συνεχόμενα μεταξύ της διάσκεψης point to point και multipoint.

Τα πρότυπα multipoint διάσκεψης πρέπει να περιλαμβάνουν τους κυβερνητικούς κανόνες απόκτησης των νέων μελών κατά τη διάρκεια μιας multipoint διάσκεψης και τις αποχωρήσεις των μελών όταν μία βίντεο-διάσκεψη είναι ακόμη σε εξέλιξη. Ο έλεγχος των διασκέψεων μεταξύ μερικών μελών ενώ η ολική βίντεο-διάσκεψη είναι σε εξέλιξη είναι επίσης ένα σημαντικό θέμα σε τέτοια περιβάλλοντα και πρέπει να διευθετηθεί με ένα σύνολο προτύπων.

- **H.261:**

Το H.261 περιγράφει τις μεθόδους κωδικοποίησης και αποκωδικοποίησης για οπτικοακουστικές υπηρεσίες σε DS1 και υπό-DS1 ρυθμούς. Το H.261 είναι μια από τις τυποποιήσεις της ITU-TS H.320 οικογένειας για βίντεο-τηλέφωνο (video phone) και τηλεδιάσκεψη σε ρυθμούς bits κυμαινόμενους από 64Kbps σε 2Mbps. Επομένως ο προτεινόμενος αλγόριθμος κωδικοποίησης video πρέπει να είναι ικανός να λειτουργήσει σε πραγματικό χρόνο με ελάχιστη καθυστέρηση. Για  $p=1$  ή  $p=2$  εξαιτίας του περιορισμένου διαθέσιμου ρυθμού bit, μόνο χαμηλής κίνησης οπτική επικοινωνία, όπως το βίντεο-τηλέφωνο, είναι κατάλληλη. Για  $p>5$ , με ένα μεγαλύτερο ρυθμό bit διαθέσιμο, πιο περίπλοκες εικόνες μπορούν να μεταφερθούν καλύτερα. Αυτό είναι συνεπώς πιο κατάλληλο για video-διάσκεψη.

Για να εξασφαλίσουμε την ενδοεπικοινωνία μεταξύ διαφορετικών συστημάτων τηλεόρασης (π.χ. PAL και NTSC) και να μειώσουμε το ρυθμό bit, το Σύνθηες Ενδιάμεσο σχήμα (CIF) υιοθετήθηκε στο H.261. Η πρόταση από το CIF είναι 288 γραμμές από 360 pixels για συνιστώσες φωτεινότητας και 144 γραμμές από 180 pixels για δυο συνιστώσες χρωματικότητας. Για χαμηλού ρυθμού bit εφαρμογές, Quarter-CIF (QCIF) χρησιμοποιείται. Η πρόταση από το QCIF είναι 144 γραμμές από 180 pixels για συνιστώσες φωτεινότητας και 72 γραμμές από 90 pixels για δυο συνιστώσες χρωματικότητας.

Έπειτα μια υβριδική κωδικοποίηση του DPCM με κίνηση και δυο διδιάστατα DCT εκτελείται στο CIF ή στο QCIF. Τελικά, κβαντοποιημένοι, μετασχηματισμένοι συντελεστές είναι μεταβλητού μήκους κωδικοποιημένοι (VLC Variable-length Coded) χρησιμοποιώντας Huffman και μεταβλητού μήκους (run-length) μεθόδους. Ο αλγόριθμος κωδικοποίησης είναι παρόμοιος με εκείνον του JPEG.

- H.320:

Το H.320 είναι το όνομα ενός πρότυπου που δημιουργήθηκε από την ITU-TS για video-τηλέφωνο και τηλεδιάσκεψη σε ρυθμούς bits κυμαινόμενους από 56 σε 1930 Kbps. Το πρότυπο αυτό έχει σχεδιαστεί για να λειτουργεί πάνω σε ψηφιακά κυκλικά δίκτυα και κυρίως κυκλικές - switched υπηρεσίες - αυτό είναι στην πράξη το ISDN - αλλά έχει επίσης εφαρμοστεί πάνω σε πακέτα δικτύων. Σε αυτό το σημείο θα πρέπει να σημειωθεί ότι το H.320 δεν είναι ένα πρότυπο ήχο-video διάσκεψης μόνο. Καλύπτει επίσης την υποστήριξη γραφημάτων και υποβαθμισμένων modes για ήχο μόνο, ή ακόμα τηλεδιάσκεψη με γραφήματα και ήχο. Επιπλέον καθορίζει τους μηχανισμούς για υποστήριξη πολυμερής τηλεδιάσκεψης.



## 5. Μετάδοση Multimedia δεδομένων

---

### 5.1 Διαδικασία δικτύου

Η μετάδοση ψηφιοποιημένων δεδομένων πολυμέσων γενικά θέτει μεγάλες απαιτήσεις στα δίκτυα[4]. Ειδικά τα διαρκή μέσα όπως είναι ο ήχος και το κινούμενο βίντεο απαιτούν υψηλό μετάδοσης, δηλαδή υψηλή χωρητικότητα μετάδοσης σε bit/sec αλλά και περιορισμένες καθυστερήσεις μετάδοσης.

### 5.2 Μετάδοση πολυμέσων

Η ανάπτυξη των οπτικών ινών στην τεχνολογία μετάδοσης - αυτό καλείται *lightware* μετάδοση - Έχει ανοίξει μία νέα εποχή για την επικοινωνία. Οι ίνες μεταδίδουν διαμορφωμένο ως αντί για διαμορφωμένα ηλεκτρικά σήματα όπως αυτά χρησιμοποιούνται στα μεταλλικά καλώδια. Με αλλά λόγια οι ίνες μεταδίδουν φωτόνια αντί για ηλεκτρόνια. Μία οπτική ίνα σήμερα μπορεί να μεταδώσει μέχρι και 2500 εκατομμύρια bits ανά δευτερόλεπτο και υποστηρίζει 32000 παράλληλες τηλεφωνικές κλήσεις. Μπορεί επίσης να υποστηρίζει χιλιάδες συμπιεσμένα ψηφιακά κανάλια TV κάθε φορά. Οι οπτικές ίνες είναι η βάση για την εξάπλωση της τεχνολογίας τοπικών δικτύων υψηλών ταχυτήτων (LAN). Στις αναπτυσσόμενες χώρες, η τηλεπικοινωνιακή υποδομή βασίζεται στις οπτικές ίνες. Αυτή η υποδομή έχει μορφοποιηθεί από τα βασικά καλώδια και τα συστήματα μετάδοσης τα οποία υποστηρίζουν όλους τους τύπους επικοινωνίας, αρχίζοντας από τη τηλεφωνία και φτάνοντας ως τις σημείο προς σημείο γραμμές.

Οι οπτικές ίνες είναι σχετικά ένα φθηνό φυσικό μέσο ειδικά για τις μακρινές αποστάσεις. Αυτό που κοστίζει πιο πολύ είναι τα συστήματα μετάδοσης τα οποία είναι προσαρτημένα κατά μήκος της ίνας, ειδικά αυτά που λειτουργούν σε υψηλή ταχύτητα. Αλλά αφότου εγκατασταθούν, μια οπτική ίνα δεν απαιτεί να εκμεταλλευθεί αμέσως όλη η χωρητικότητα. Πιο αργός και φθηνός εξοπλισμός μετάδοσης μπορεί αρχικά να χρησιμοποιηθεί, ενώ σχεδόν απεριόριστη χωρητικότητα είναι ενδεχομένως διαθέσιμη για μελλοντικές αναπτύξεις.

### 5.3 Πρότυπα για μετάδοση βίντεο

- **H.261:**

Το βασικό πρότυπο για βίντεο διάσκεψη και δια-πρακτικότητα των multimedia είναι ο H.261[1], ο οποίος καθορίζει διάφορες διαδικασίες για codecs ώστε επικοινωνεί σε ρυθμούς πλαισίου μέχρι 30 fps μέσα σε δίκτυα που αποτελούνται από πολλαπλές ευκολίες με εύρος ζώνης 64-Kbps.

Η CCITT υιοθέτησε τον τύπο H.261, γνωστό επίσης ως p\*64, ειδικά σχεδιασμένο να επιτύχει μεταφορά εικόνων βίντεο μέσα από ψηφιακά δίκτυα σε ρυθμούς δεδομένων που κυμαίνονται από 64 σε 2048 Kbps. Μέχρι τότε οι πωλητές βίντεο διάσκεψης χρησιμοποιούσαν ιδιοκτησιακές τυποποιήσεις με διάφορες τεχνικές συμπίεσης οι οποίες ήταν ασύμβατες μεταξύ τους και βελτιστοποιημένες για βίντεο διάσκεψης μεταδόσεις στα ιδιωτικά δίκτυα. Το H.261 περιέχει ένα τύπο συμπίεσης που βασίζεται στην τεχνική DCT και είναι παρόμοιο με το MPEG αλλά επιτρέπει πού μεγαλύτερες ταχύτητες μετάδοσης. Το H.261 «θυσιάζει» την ποιότητα βίντεο και το μέγεθος της εικόνας ώστε παράσχει βίντεο διάσκεψης ενδομπρακτικότητα στις τυποποιήσεις QCIF και CIF σε χαμηλότερο εύρος ζώνης 112- και 128-Kbps.

Πρωταρχικά αυτό το πρότυπο σχεδιάστηκε για έλεγχο από σημείο σε σημείο βίντεο διάσκεψης, αλλά η χρήση αυτής της τεχνολογίας τώρα επικαλύπτεται από τις εφαρμογές multimedia. Με την πρόοδο καινούριες λειτουργίες ήρθαν στο προσκήνιο οι οποίες απαιτούν παραπάνω πρότυπα μετάδοσης. Έτσι το H.261 αναπτύσσεται για να καλύψει έναν αριθμό από αποφασιστικές λειτουργίες.

- **H.233:**

Το πρότυπο αυτό καθορίζει μια μέθοδο για χρήση μιας μοναδικής απόκρυψης μεταξύ εξουσιοδοτημένων βίντεο διασκέψεων μερών πριν κάποια συνεργατική ενέργεια μπορεί να συμβεί. Δεν είναι ακόμα τελείως κατανοητό ποια θέματα πληροφορίας και δεδομένων ασφάλειας εμπλέκονται στην πολυμερή σύσκεψη των multimedia. Αυτά, μαζί με πρόσβαση από πολλούς χρήστες για συνεταιρισμό βάσεων δεδομένων, ευπαθή πληροφορία μάρκετινγκ και προσωπικό καταλόγων πρέπει να προστατευθεί για αξιοπιστία σε διάφορα επίπεδα της multipoint διάσκεψης.

Τέτοια θέματα είναι σημαντικά στις επικοινωνίες των multimedia στον ιατρικό χώρο και πρέπει να λυθούν πριν αυτά τα συστήματα χρησιμοποιηθούν ευρέως. Υπάρχουν πολλές εικόνες και βίντεο σε τέτοια ιατρικά περιβάλλοντα που περιέχουν υψηλά ευπαθή πληροφορίες σχετικές με ασθενείς, όπως την υγεία τους, ασθένειες, ιατρικές θεραπείες, οικονομικής κατάστασης και συνθήκες σχετικές με το σώμα και την ψυχή. Το πρότυπο απόκρυψης μπορεί να καταλήξει να είναι μια από τις πιο σημαντικές εξελίξεις οι οποίες θα λάβουν μέρος πριν οι ουσιαστικές συνεταιριστικές δραστηριότητες εξελιχθούν.

- **H.320**

Το H.320 είναι η προστασία CCITT προτύπου συμμόρφωσης που καθορίζει όλες τις τεχνικές απαιτήσεις για μεταφορά στενού εύρους ζώνης των ήχο-οπτικών συστημάτων.

Το πρότυπο αυτό υποδηλώνει συμμόρφωση με τον βασικό τύπο κωδικοποίησης H.261, έναν αριθμό βοηθητικών τύπων και ένα νέο τύπο ηχητικής συμπίεσης. Αυτοί οι συνιστάμενοι τύποι περιλαμβάνουν τον H.221, ο οποίος ασχολείται με πλαισιωμένη πληροφορία. Ο H.230, για χειρισμό σημάτων ελέγχου και ένδειξης και ο H.242 ο οποίος ελέγχει την μετάδοση, εγκαθιστά και αποσυνδέει. Το νέο πρότυπο G.728, ο οποίος ασχολείται με την συμπίεση ήχου στα 16 Kbps, είναι επίσης μέλος αυτής της οικογένειας. Ο παρακάτω πίνακας κάνει περίληψη όλης της οικογένειας τύπων H.320 υιοθετημένων και σε εξέλιξη.

Όλοι οι τύποι H - Έχουν υιοθετηθεί το 1990, ενώ ο G.728 εξεδόθη το 1992. Είναι ξεκάθαρο ότι αυτή η οικογένεια τύπων ακόμα εξελίσσεται και τα πρωτόκολλα multimedia και οι τύποι MPEG μπορεί να γίνουν μέρος του H.320. Ο H.320 είναι διαθέσιμος σε ένα απλό microchip την Ενοποιημένη Τεχνολογία Πληροφορία η οποία επίσης περιλαμβάνει JPEG, MPEG και άλλους αλγόριθμους. Η AT&T ανέπτυξε επίσης το σύνολο των μικροτσιπ AVP το οποίο παρουσιάζει την ίδια λειτουργία.

- **H.263**

Η σύσταση H.263 σχεδιάστηκε για επικοινωνία σε χαμηλούς ρυθμούς δεδομένων. Η σύσταση H.263 χρησιμοποιεί μία block motion-compensated DCT μέθοδο για την κωδικοποίηση του video. Η κωδικοποίηση κατά H.263 είναι πολύ πιο αποτελεσματική από ότι η κωδικοποίηση κατά H.261. Στην σύσταση H.263, η κωδικοποίηση πραγματοποιείται τεμαχίζοντας κάθε εικόνα σε macroblocks. Κάθε macroblock αποτελείται από 16x16 luminance blocks και 8x8 chrominance blocks. Κάθε macroblock μπορεί να κωδικοποιηθεί είτε ως intra (χωρίς συσχετίσεις ανάμεσα σε διαδοχικά frames) είτε ως inter (με συσχετίσεις ανάμεσα σε διαδοχικά frames). Η τεχνική DCT η οποία χρησιμοποιείται, παρέχει spatial redundancy και temporal redundancy. Η σύσταση H.263 στηρίζεται στην H.261 και παρέχει κάποιες επεκτάσεις για την υποστήριξη πιο αποτελεσματικής κωδικοποίησης.

## ΜΕΡΟΣ Β: Πρωτόκολλα πραγματικού χρόνου

### 1.ΕΙΣΑΓΩΓΗ

---

#### 1.1 ΕΞΕΛΙΞΗ ΤΟΥ ΔΙΚΤΥΟΥ

##### 1.1.1 Δεκαετία του 60: Η εκκίνηση

Το σημερινό Internet αποτελεί εξέλιξη του ARPANET, ενός δικτύου που άρχισε να αναπτύσσεται πειραματικά στα τέλη της δεκαετίας του 60 στις Ηνωμένες Πολιτείες[6]. Στα πανεπιστήμια των ΗΠΑ οι ερευνητές ξεκινούν να πειραματίζονται με τη διασύνδεση απομακρυσμένων υπολογιστών μεταξύ τους. Το δίκτυο ARPANET γεννιέται το 1969 με πόρους του προγράμματος ARPA (Advanced Research Project Agency) του Υπουργείου Άμυνας, με σκοπό να συνδέσει το Υπουργείο με στρατιωτικούς ερευνητικούς οργανισμούς και να αποτελέσει ένα πείραμα για τη μελέτη της αξιόπιστης λειτουργίας των δικτύων. Στην αρχική του μορφή, το πρόγραμμα απέβλεπε στον πειραματισμό με μια νέα τεχνολογία γνωστή σαν μεταγωγή πακέτων (packet switching), σύμφωνα με την οποία τα προς μετάδοση δεδομένα κόβονται σε πακέτα και πολλοί χρήστες μπορούν να μοιραστούν την ίδιο επικοινωνιακή γραμμή.

Στόχος ήταν η δημιουργία ενός διαδικτύου που θα εξασφάλιζε την επικοινωνία μεταξύ απομακρυσμένων δικτύων, έστω και αν κάποια από τα ενδιάμεσα συστήματα βρίσκονταν προσωρινά εκτός λειτουργίας. Κάθε πακέτο θα είχε την πληροφορία που χρειαζόνταν για να φτάσει στον προορισμό του, όπου και θα γινόταν η επανασύνθεσή του σε δεδομένα τα οποία μπορούσε να χρησιμοποιήσει ο τελικός χρήστης. Το παραπάνω σύστημα θα επέτρεπε σε υπολογιστές να μοιράζονται δεδομένα και σε ερευνητές να υλοποιήσουν το ηλεκτρονικό ταχυδρομείο.

##### 1.1.2 Δεκαετία του 70: οι πρώτες συνδέσεις

Το 1973, ξεκινά ένα νέο ερευνητικό πρόγραμμα που ονομάζεται Interneting Project (Πρόγραμμα Διαδικτύωσης) προκειμένου να ξεπεραστούν οι διαφορετικοί τρόποι που χρησιμοποιεί κάθε δίκτυο για να διακινεί τα δεδομένα του. Στόχος είναι η διασύνδεση πιθανώς ανόμοιων δικτύων και η ομοιόμορφη διακίνηση δεδομένων από το ένα δίκτυο στο άλλο[6]. Από την έρευνα γεννιέται μια νέα τεχνική, το Internet Protocol (IP) (Πρωτόκολλο Διαδικτύωσης), από την οποία θα πάρει αργότερα το όνομα του το Internet. Διαφορετικά δίκτυα που χρησιμοποιούν το κοινό πρωτόκολλο IP μπορούν να συνδέονται και να αποτελούν ένα διαδίκτυο.

Σε ένα δίκτυο IP όλοι οι υπολογιστές είναι ισοδύναμοι, οπότε τελικά οποιοσδήποτε υπολογιστής του υπολογιστή του διαδικτύου μπορεί να επικοινωνεί με οποιονδήποτε άλλον. Επίσης σχεδιάζεται μια άλλη τεχνική για τον έλεγχο της μετάδοσης των δεδομένων, το Transmission Control Protocol (TCP) (Πρωτόκολλο Ελέγχου Μετάδοσης). Ορίζονται προδιαγραφές για τη μεταφορά αρχείων μεταξύ υπολογιστών (FTP) και για το ηλεκτρονικό ταχυδρομείο (E-mail). Σταδιακά συνδέονται με το ARPANET ιδρύματα από άλλες χώρες, με πρώτα το University College of London (Αγγλία) και το Royal Radar Establishment (Νορβηγία).

### **1.1.3 Δεκαετία του 80: ένα παγκόσμιο δίκτυο για την ακαδημαϊκή κοινότητα**

Το 1983, το πρωτόκολλο TCP/IP (δηλ. ο συνδυασμός των TCP και IP) αναγνωρίζεται ως πρότυπο από το Υπουργείο Άμυνας των ΗΠΑ. Η έκδοση του λειτουργικού συστήματος Berkeley UNIX το οποίο περιλαμβάνει το Εκατοντάδες Πανεπιστήμια συνδέουν τους υπολογιστές τους στο ARPANET, το οποίο επιβαρύνεται πολύ και το 1983, χωρίζεται σε δύο τμήματα: στο MILNET (για στρατιωτικές επικοινωνίες) και στο νέο ARPANET (για χρήση αποκλειστικά από την πανεπιστημιακή κοινότητα και συνέχιση της έρευνας στη δικτύωση). Το 1985, το National Science Foundation (NSF) δημιουργεί ένα δικό του γρήγορο δίκτυο, το NSFNET χρησιμοποιώντας το πρωτόκολλο TCP/IP, προκειμένου να συνδέσει πέντε κέντρα υπερ-υπολογιστών μεταξύ τους και με την υπόλοιπη επιστημονική κοινότητα. Στα τέλη της δεκαετίας του '80, όλο και περισσότερες χώρες συνδέονται στο NSFNET (Καναδάς, Γαλλία, Σουηδία, Αυστραλία, Γερμανία, Ιταλία, κ.α.) Χιλιάδες πανεπιστήμια και οργανισμοί δημιουργούν τα δικά τους δίκτυα και τα συνδέουν πάνω στο παγκόσμιο αυτό δίκτυο το οποίο αρχίζει να γίνεται γνωστό σαν INTERNET και να εξαπλώνεται με τρομερούς ρυθμούς σε ολόκληρο τον κόσμο. Το 1990, το ARPANET πλέον καταργείται.

### **1.1.4 Δεκαετία 90: ένα παγκόσμιο δίκτυο για όλους**

Όλο και περισσότερες χώρες συνδέονται στο NSFNET, μεταξύ των οποίων και η Ελλάδα το 1990. Το 1993, το εργαστήριο CERN στην Ελβετία παρουσιάζει το World Wide Web (WWW) (παγκόσμιο Ιστό) που αναπτύχθηκε από τον Tim Berners-Lee. Πρόκειται για ένα σύστημα διασύνδεσης πληροφοριών σε μορφή πολυμέσων (multimedia) που βρίσκονται αποθηκευμένες σε χιλιάδες υπολογιστές του Internet σε ολόκληρο τον κόσμο και παρουσίασης τους σε ηλεκτρονικές σελίδες, στις οποίες μπορεί να περιηγηθεί κανείς χρησιμοποιώντας το ποντίκι. Το γραφικό αυτό περιβάλλον έκανε την εξερεύνηση του Internet προσιτή στον απλό χρήστη.

Παράλληλα, εμφανίζονται στο Internet διάφορα εμπορικά δίκτυα που ανήκουν σε εταιρίες παροχή υπηρεσιών Internet (Internet Service Providers - ISP) και προσφέρουν πρόσβαση στο Internet για όλους. Οποιοσδήποτε διαθέτει PC και modem μπορεί να συνδεθεί με το Internet σε τιμές που μειώνονται διαρκώς. Το 1995, το NSFNET καταργείται πλέον επίσημα και το φορτίο του μεταφέρεται σε εμπορικά δίκτυα.

## **2. Multimedia Over IP:**

---

### **2.1 Multimedia networking: Στόχοι και προοπτικές**

Τα δίκτυα υπολογιστών σχεδιάστηκαν για να συνδέονται υπολογιστές που βρίσκονται σε διαφορετικές τοποθεσίες και να ανταλλάσσουν δεδομένα. Παλιότερα, όλα τα δεδομένα που μεταφέρονταν μέσω δικτύου ήταν κείμενα (textual)[9]. Σήμερα, με την ανάπτυξη των πολυμεσικών και δικτυακών τεχνολογιών, η πολυμεσική πληροφορία έχει κυριαρχήσει στο Internet. Το animation, ο ήχος και τα video clips γίνονται όλο και πιο δημοφιλή.

Multimedia networking σημαίνει να χτίσουμε το hardware, το software και τα εργαλεία των εφαρμογών με τέτοιο τρόπο ώστε οι χρήστες να μπορούν να επικοινωνούν με πολυμεσική πληροφορία. Η μετάδοση πολυμεσικής πληροφορίας θα μεταβάλλει τον υπολογιστή σε ένα επικοινωνιακό εργαλείο.

### **2.2 Επικοινωνία πραγματικού χρόνου**

Το multimedia networking δεν είναι απλή διαδικασία. Μερικά από τα προβλήματα που παρουσιάζονται είναι τα παρακάτω:

Σε σύγκριση με τις παραδοσιακές εφαρμογές που υποστηρίζουν μόνο κείμενο (textual applications), οι πολυμεσικές εφαρμογές απαιτούν συνήθως πολύ μεγαλύτερη χωρητικότητα (bandwidth).

Οι περισσότερες πολυμεσικές εφαρμογές απαιτούν επικοινωνία σε πραγματικό χρόνο. Στη μετάδοση του ήχου και της εικόνας πρέπει να υπάρχει συνέχεια (να μη διακόπτεται η μετάδοση). Αν τα δεδομένα δε φτάνουν στην ώρα τους για να μεταδοθούν, δημιουργούνται κενά που γίνονται αντιληπτά από το χρήστη και υποβαθμίζουν την ποιότητα της (πολυμεσικής) πληροφορίας. Στη διαδικτυακή τηλεφωνία, μια καθυστέρηση μέχρι 250 ms θεωρείται ανεκτή. Αν η καθυστέρηση ξεπεράσει την τιμή αυτή, η ποιότητα της επικοινωνίας είναι χαμηλή. Ο φόρτος του δικτύου παίζει επίσης σημαντικό ρόλο στην επικοινωνία πραγματικού χρόνου. Αν υπερφορτωθεί το δίκτυο τα δεδομένα πραγματικού χρόνου δε μπορούν να φτάσουν στην ώρα τους.

Στην περίπτωση που ξαναστέλνονται τα πακέτα που δεν έφτασαν στον προορισμό τους, το πρόβλημα γίνεται εντονότερο, και μπορεί να οδηγήσει μέχρι και στην κατάρρευση του δικτύου.

Η ροή πολυμεσικών δεδομένων (multimedia data stream), είναι συνήθως bursty. Το να αυξήσουμε τη χωρητικότητα δε θα λύσει το πρόβλημα αυτό. Για τις περισσότερες πολυμεσικές εφαρμογές, ο παραλήπτης έχει ένα περιορισμένου μεγέθους buffer. Αν δε ληφθούν μέτρα για την εξομάλυνση του data stream, μπορεί να έχουμε υπερχειλίση του buffer της εφαρμογής. Όταν τα δεδομένα έρχονται πολύ γρήγορα, ο buffer μπορεί να υπερχειλίσει και να χαθούν μερικά πακέτα με αποτέλεσμα να υποβαθμιστεί η ποιότητα. Όταν πάλι τα δεδομένα έρχονται σε πού αργούς ρυθμούς η εφαρμογή υπολειτουργεί και η ποιότητα είναι προφανώς χαμηλή.

Το πρόβλημα με τα δίκτυα είναι ότι συνδέουν χιλιάδες και εκατομμύρια δέκτες, έχουν περιορισμένη χωρητικότητα, καθυστερήσεις και φόρτο που δε μπορούν να προβλεφθούν.

Η αντιμετώπιση των προβλημάτων αυτών είναι μια πρόκληση η οποία πρέπει να αντιμετωπιστεί με βάση τις υπάρχουσες τεχνολογίες τόσο σε software όσο και σε hardware. Η βάση του Internet (TCP/IP) και UDP/IP παρέχει ένα σύνολο υπηρεσιών που μπορούν να χρησιμοποιήσουν οι πολυμεσικές εφαρμογές. Γρήγορα δίκτυα (Gigabit Ethernet, ATM) παρέχουν τις υψηλές χωρητικότητες που απαιτούνται για τη μετάδοση video και ήχου.

### **2.3 Multimedia στο Internet**

Υπάρχουν και άλλοι τρόποι για τη μετάδοση πολυμεσικών δεδομένων (μισθωμένες γραμμές, ATM). Παρόλα αυτά, η ιδέα της διακίνησης πολυμεσικής πληροφορίας μέσω του Internet είναι εξαιρετικά ελκυστική αφού, το Internet εξαπλώνεται ταχύτατα και έχει εξελιχθεί στην πλατφόρμα των περισσότερων δικτυακών δραστηριοτήτων. Αυτός είναι ο κύριος λόγος για την ανάπτυξη πρωτοκόλλων για τη μετάδοση πολυμεσικής πληροφορίας πάνω από το Internet.

Σήμερα, τα πρωτόκολλα IP και Ethernet είναι τα περισσότερο διαδεδομένα για desktops και LANs, ενώ το ATM χρησιμοποιείται κυρίως στα wide area networks.

Το Internet δεν είναι κατασκευασμένο για real time επικοινωνία και συνεπώς η μετάδοση πολυμεσικής πληροφορίας συναντάει προβλήματα που πρέπει να λυθούν.

**Πρώτον**, η απαιτούμενη πληροφορία είναι μεγάλη σε όγκο. Συνεπώς το hardware θα πρέπει να παρέχει αρκετή χωρητικότητα.

**Δεύτερον**, συνήθως τα πολυμεσικά δεδομένα δε στέλνονται σε ένα χρήστη, αλλά σε μία ομάδα χρηστών. Στους χρήστες στέλνεται ένα data stream και όχι πολλαπλά αντίγραφα (multicast) και αυτό θα πρέπει να προβλέπεται από το πρωτόκολλο.

**Τρίτον**, οι εφαρμογές πραγματικού χρόνου απαιτούν εγγυημένη χωρητικότητα κατά τη διάρκεια μετάδοσης των δεδομένων, και αυτό δεν ισχύει στο Internet. Συνεπώς θα πρέπει να υπάρχουν κάποιοι μηχανισμοί για τις εφαρμογές πραγματικού χρόνου, οι οποίοι θα δεσμεύουν πόρους κατά μήκος του μονοπατιού μετάδοσης.

**Τέταρτον**, το Internet είναι ένα packet-switching datagram δίκτυο όπου τα πακέτα δρομολογούνται ανεξάρτητα το ένα από το άλλο. Το γεγονός αυτό είναι δυνατό να εισαγάγει ανεπιθύμητες καθυστερήσεις. Χρειάζονται συνεπώς ειδικά πρωτόκολλα που να λαμβάνουν υπόψη τους χρονικούς περιορισμούς.

**Πέμπτον**, είναι απαραίτητη η ύπαρξη κάποιων λειτουργιών για το χειρισμό της παρουσίασης των πολυμεσικών πληροφοριών.

Οι απαντήσεις στα παραπάνω προβλήματα είναι τα πρωτόκολλα που θα συζητηθούν στην εργασία αυτή.

### **3. Γιατί τα παραδοσιακά πρωτόκολλα μεταφοράς δεν είναι κατάλληλα για multimedia networking**

---

Τα παραδοσιακά πρωτόκολλα μεταφοράς (TCP / UDP) έχουν σχεδιαστεί για αξιόπιστες επικοινωνίες δεδομένων σε δίκτυα με χαμηλό εύρος ζώνης και υψηλό σφαλμάτων[7]. Συνεπώς δεν είναι βελτιστοποιημένα για λειτουργία σε υψηλές ταχύτητες. Επίσης, δεν υποστηρίζουν QOS και multicast. Συμπεραίνουμε λοιπόν ότι είναι ακατάλληλα για επικοινωνίες πολυμέσων. Παρακάτω περιγράφουμε ορισμένα χαρακτηριστικά των πρωτοκόλλων αυτών που είναι ακατάλληλα για υψηλής ταχύτητας εφαρμογές πολυμέσων.

#### **3.1 Δομή των πρωτοκόλλων**

Οι δύο κύριοι στόχοι των συνόλων πρωτοκόλλων OSI και TCP/IP ήταν να σπάσουν το πρόβλημα της κατασκευής δικτύου σε απλά διαχειρίσιμα στρώματα ή επίπεδα λειτουργιών και να επιτρέχουν συνεργασία ανάμεσα στα διαφορετικά επίπεδα λειτουργικότητας των πρωτοκόλλων δικτύου. Η υψηλή απόδοση ξεκάθαρα δεν ήταν ανάμεσα στους στόχους των πρωτοκόλλων με στρώματα.

Η φύση των πρωτοκόλλων που είναι υλοποιημένα με στρώματα εισάγει μια έμφυτη συμφόρηση στην επικοινωνία. Τα ακατέργαστα δεδομένα μετακινούνται από ένα στρώμα στο άλλο και σε κάθε στρώμα υφίστανται επεξεργασία ξανά. Στην τυπική περίπτωση αντιγράφονται σε ένα νέο χώρο διευθύνσεων στη μνήμη και ένα νέο header προστίθενται σε αυτά.



Το πακέτο δεδομένων αποθηκεύεται και προωθείται από κάθε στρώμα. Αυτή η περιττή αντιγραφή συνεισφέρει στην καθυστέρηση του δικτύου. Η αντιγραφή δεδομένων είναι χρονοβόρα επειδή η προσπέλαση μνήμης είναι μια αργή διαδικασία συγκρινόμενη με την ταχύτητα της κεντρικής μονάδας επεξεργασίας, και κάθε byte δεδομένων πρέπει να μετακινείται κάθε φορά που γίνεται αντιγραφή δεδομένων.

Η λύση σε αυτό το πρόβλημα είναι να μειώσουμε τον όγκο των δεδομένων που αντιγράφονται και να περνάμε δείκτες ανάμεσα στα στρώματα αντί να αντιγράφουμε κάθε φορά τα ίδια τα δεδομένα.

### 3.2 Πολύπλεξη

Η πολύπλεξη είναι μια λειτουργία σύμφωνα με την οποία stream πολλών εφαρμογών πολυπλέκονται με σκοπό να περάσουν μέσα από μια μόνο σύνδεση. Το κύριο αποτέλεσμα της πολύπλεξης είναι το γεγονός ότι οι ανώτεροι πελάτες μοιράζονται τους πόρους ενός χαμηλότερου επιπέδου. Μερικές φορές η πολύπλεξη είναι απαραίτητη, όπως στην περίπτωση που πολλές συνδέσεις μοιράζονται το ίδιο μέσω μετάδοσης. Εντούτοις η πολύπλεξη σε μεγάλο αριθμό επιπέδων μειώνει την απόδοση και αυτό γιατί: Η πολύπλεξη καθώς και η αντίστροφη λειτουργία αυξάνουν την πολυπλοκότητα των πρωτοκόλλων και των υλοποιήσεών τους, καταλήγοντας τελικά να προσφέρουν χαμηλό throughput.

Οι ροές δεδομένων σε διαφορετικές συνδέσεις επηρεάζουν η μία την άλλη, προκαλώντας καθυστέρηση.

Συνδέσεις που δυσλειτουργούν δεν μπορούν να αναγνωριστούν και να αντιμετωπιστούν στο κατώτερο στρώμα, προκαλώντας έτσι καθυστέρηση στα δεδομένα άλλων συνδέσεων.

Διαφορετικές συνδέσεις με διαφορετικές QOS απαιτήσεις αντιμετωπίζονται με τον ίδιο τρόπο, με αποτέλεσμα να έχουμε είτε σπατάλη πόρων είτε αδυναμία να ικανοποιήσουμε όλες τις QOS απαιτήσεις.

Η λύση σε αυτά τα προβλήματα είναι να χρησιμοποιήσουμε διαφορετικά threads για διαφορετικές ροές ή να χρησιμοποιήσουμε διαφορετικές στοίβες μεταφοράς για διαφορετικές ροές.

### 3.3 Έλεγχος ροής

Ο παραδοσιακός μηχανισμός ελέγχου ροής είναι ο έλεγχος ροής ολισθαίνοντος παραθύρου, ο οποίος επιτρέπει να μεταδοθεί ένας σταθερός αριθμός από bytes (ένα παράθυρο από bytes), χωρίς να χρειάζεται επιβεβαίωση ή βεβαίωση λήψης από τον παραλήπτη. Μετά από τη μετάδοση αυτού του σταθερού αριθμού από bytes, ο πομπός μπορεί να μεταδώσει περισσότερα δεδομένα μόνο όταν του το επιτρέψει ο παραλήπτης στέλλοντας μία βεβαίωση λήψης.

Το μέγεθος παραθύρου που χρησιμοποιεί το TCP είναι 64 kbytes. Για ένα αργό δίκτυα, αυτό το μέγεθος παράθυρο θεωρείται πολύ μεγάλο. Για παράδειγμα, αν η ταχύτητα μετάδοσης είναι 64 kbps, χρειάζονται 8 δευτερόλεπτα για να μεταδοθούν 64 kbytes δεδομένων. Ο συνηθισμένος round trip χρόνος του δικτύου είναι αρκετά μικρότερος των 8 δευτερολέπτων, οπότε πριν ο πομπός τελειώσει να μεταδίδει τα δεδομένα του τρέχοντος παραθύρου, λαμβάνει μια βεβαίωση λήψης για το αν η ενδιάμεση μνήμη (buffer) του παραλήπτη δεν είναι γεμάτη έτσι ώστε ο πομπός να μπορεί να μεταδίδει δεδομένα συνεχώς.

Για μετάδοση υψηλής ταχύτητας ο έλεγχος ροής ολισθαινόντος παραθύρου δεν είναι κατάλληλος και αυτό γιατί:

Το μέγεθος παραθύρου είναι υπερβολικά μικρό και ο πομπός την περισσότερη ώρα θα περιμένει για την άδεια μετάδοσης από τον παραλήπτη. Έτσι, το εύρος ζώνης της μετάδοσης δε χρησιμοποιείται πλήρως. Για παράδειγμα, ένας πομπός θα στείλει 64 kbytes μέσα σε 50 ms με μια ταχύτητα των 10 Mbps. Για ένα WAN, ο round trip χρόνος είναι συνήθως πολύ μεγαλύτερος από 50 ms. Μια μερική λύση για το πρόβλημα αυτό είναι η χρησιμοποίηση ενός μεγαλύτερου μεγέθους παραθύρου.

Ο έλεγχος ροής ολισθαινόντος παραθύρου από μόνος του δεν είναι κατάλληλος για δεδομένα πολυμέσων. Ο έλεγχος ροής ολισθαινόντος παραθύρου υποθέτει ότι ο ρυθμός μετάδοσης bits μπορεί να προσαρμόζεται ανάλογα με την κατάσταση του δικτύου και του παραλήπτη. Αυτό δεν είναι δυνατό για συνεχή μέσα, τα οποία πρέπει να στέλνονται με τον εσωτερικό τους ρυθμό δεδομένων. Διακοπτόμενη μετάδοση και λήψη δεδομένων ήχου και εικόνας είναι άχρηστη για εφαρμογές πολυμέσων πραγματικού χρόνου.

### 3.4 Έλεγχος λαθών

Τα TCP παρέχει αξιόπιστη επικοινωνία δεδομένων. Όταν κάποιο πακέτο χάνεται ή αλλοιώνεται, τότε το πακέτο αυτό μεταδίδεται ξανά. Αυτή η στρατηγική δεν είναι κατάλληλη για επικοινωνίες πολυμέσων και αυτό γιατί:

- Στα δεδομένα πολυμέσων είναι ανεκτά κάποια λάθη ή απώλειες.
- Η επαναμετάδοση προκαλεί καθυστέρηση στα επόμενα δεδομένα με αποτέλεσμα οι παραλήπτες να λαμβάνει περισσότερα άχρηστα δεδομένα (μετέπειτα δεδομένα είναι άχρηστα 'σο και τα χαμένα δεδομένα όταν πρόκειται για συνεχή μέσα).
- Η υλοποίηση μηχανισμού επαναμετάδοσης απαιτεί κάποιον αριθμό χρονομέτρων (timers) και μεγάλες ενδιάμεσες μνήμες (buffers), καθιστώντας έτσι το πρωτόκολλο μεταφοράς πολύπλοκο και αργό.

Για επικοινωνίες πολυμέσων, πρέπει να παρέχεται ανίχνευση λαθών και να αποφασίζει η εφαρμογή αν χρειάζεται επαναμετάδοση ή όχι. Όταν απαιτείται επαναμετάδοση, μια επιλεκτική επαναμετάδοση είναι προτιμότερη από μια στρατηγική N-οπισθοδρόμησης. Στην επιλεκτική επαναμετάδοση, μόνο τα χαμένα πακέτα ή τα πακέτα με λάθη επαναμεταδίδονται. Σε μια τεχνική N-οπισθοδρόμησης, επαναμεταδίδονται όλα τα πακέτα από το τελευταίο λάθος ή απώλεια, παρόλο που τα περισσότερα πακέτα έχουν φτάσει σωστά στον προορισμό τους.

Μια άλλη λύση είναι να χρησιμοποιήσουμε κώδικα διόρθωσης σφαλμάτων, σύμφωνα με τον οποίο στέλνονται και άλλες πληροφορίες ώστε να καθίσταται δυνατή η διόρθωση σφαλμάτων στον παραλήπτη χωρίς να υπάρχει ανάγκη για επαναμετάδοση. Το μειονέκτημα της παραπάνω τεχνικής είναι ότι καταναλώνει πρόσθετο εύρος ζώνης.

### 3.5 Πληροφορίες ελέγχου

Όλα τα πρωτόκολλα λειτουργούν ανταλλάσσοντας πληροφορίες κατάστασης πρωτοκόλλου, είτε με το να επισυνάπτουν αυτές τις πληροφορίες σε ένα πακέτο με headers είτε με το να στέλνουν ειδικά πακέτα ελέγχου πρωτοκόλλου που δεν περιέχουν δεδομένα είτε και τα δύο. Σε κάθε περίπτωση, η μηχανή κατάστασης πρωτοκόλλου πρέπει να αναλύσει τις πληροφορίες ελέγχου πρωτοκόλλου για να αποκωδικοποιήσει τα περιεχόμενά τους.

Αρκετοί παράγοντες επηρεάζουν την πολυπλοκότητα της υλοποίησης και το throughput του πρωτοκόλλου:

Ο πρώτος παράγοντας είναι αν η θέση της πληροφορίας ελέγχου είναι σταθερή μέσα στο κάθε πακέτο. Η σταθερή θέση επιτρέπει πιο απλή υλοποίηση και υψηλότερο throughput.

Ο δεύτερος παράγοντας είναι αν η πληροφορία ελέγχου είναι ευθυγραμμισμένη με τα bytes της μηχανής ή τα όρια των λέξεων (words). Η ευθυγράμμιση με τα bytes ή τις λέξεις επιτρέπει γρηγορότερη υλοποίηση.

### 3.6 Έλλειψη χαρακτηριστικών

Τα παραπάνω θέματα ασχολούνται κυρίως με την αποδοτικότητα και το throughput των πρωτοκόλλων μεταφοράς. Αυτά τα θέματα μπορούν να αντιμετωπιστούν ως ένα βαθμό και πρωτοκολλά μεταφοράς υψηλής ταχύτητας μπορούν να επιτευχθούν με βελτιστοποίηση, παράλληλα με υλοποίηση στο υλικό. Πραγματικά, αναφέρεται ότι μια ταχύτητα που είναι κοντά σε gigabits ανά δευτερόλεπτο μπορεί να επιτευχθεί βελτιστοποιώντας την υλοποίηση του TCP. Όπως έχει σημειωθεί νωρίτερα, η υψηλή ταχύτητα είναι απαραίτητη, αλλά άλλα χαρακτηριστικά, QOS εγγύηση και multicast, απαιτούνται για επικοινωνίες πολυμέσων. Στο TCP, δεν υπάρχει η έννοια του QOS.

#### 4. Χαρακτηριστικά που πρέπει να έχουν τα πρωτόκολλα πραγματικού χρόνου

---

Η λύση για multimedia over IP[8], είναι να κατηγοριοποιήσουμε τις εφαρμογές, να δώσουμε προτεραιότητες στις διάφορες κατηγορίες και να δεσμεύουμε τους πόρους του δικτύου με βάση τις προτεραιότητες αυτές. Το Integrated Services working group στο IETF (Internet Engineering Task Force), ανέπτυξε ένα εξελιγμένο μοντέλο υπηρεσιών Internet που καλείται Integrated Services και περιλαμβάνει best-efford service και real-time service. Το Resource ReSerVation Protocol (RSVP), μαζί με το Real Time Transport Protocol (RTP), το Real Time Control Protocol (RTCP) και το Real Time Streaming Protocol (RTSP), παρέχουν τη βάση για υπηρεσίες πραγματικού χρόνου.

Ένα πρωτόκολλο μεταφοράς πολυμέσων διαφέρει από ένα συμβατικό πρωτόκολλο μεταφοράς δεδομένων στο ότι θα πρέπει να παρέχει QOS εγγυήσεις σε εφαρμογές πολυμέσων. Γνωρίζουμε ότι, υπό κατάλληλες προϋπόθεσης, είναι δυνατό να υποστηριχτούν οι επικοινωνίες πολυμέσων στο επίπεδο δικτύου. Η λειτουργία του πρωτοκόλλου μεταφοράς είναι να εγκαθιστά και να συντηρεί με σύνδεση με χαρακτηριστικά QOS για χρήση από τις εφαρμογές. Τρεις κύριες απαιτήσεις που πρέπει να έχει ένα πρωτόκολλο πραγματικού χρόνου είναι ότι το throughput πρέπει να είναι υψηλό, ότι το πρωτόκολλο πρέπει να υποστηρίζει multicast, και ότι ένα πρωτόκολλο πραγματικού χρόνου θα πρέπει να έχει χαρακτηριστικά QOS.

##### 4.1 Υψηλό throughput

Τα δεδομένα πολυμέσων, και ειδικά το video, απαιτούν συνεχές εύρος ζώνης υψηλής μετάδοσης. Για παράδειγμα, ένα συμπεσμένο video υψηλής ποιότητας απαιτεί ένα εύρος ζώνης των 5 Mbps περίπου. Ένα μη συμπεσμένο video απαιτεί ένα εύρος ζώνης 50 με 100 φορές μεγαλύτερο από το παραπάνω[1]. Οπότε το πρωτόκολλο μεταφοράς θα πρέπει να είναι αρκετά γρήγορο ώστε να υποστηρίζει τις απαιτήσεις της εφαρμογής σε εύρος ζώνης. Εφόσον μια εφαρμογή μπορεί να χρειάζεται έναν αριθμό από συνέχεια ροές δεδομένων, η ταχύτητα του πρωτοκόλλου μεταφοράς πρέπει να είναι μεγαλύτερη από τις απαιτήσεις των παραπάνω ροών σε αθροιστικό εύρος ζώνης.

Ένας άλλος τρόπος να δούμε την απαιτήσεις ενός πρωτοκόλλου μεταφοράς σε throughput είναι από τη σκοπιά του συνολικού συστήματος επικοινωνιών. Το throughput ενός πρωτοκόλλου μεταφοράς πρέπει να είναι υψηλότερο από την ταχύτητα προσπέλασης του δικτύου. Διαφορετικά, το εύρος ζώνης που παρέχεται από τα σημεία προσπέλασης του δικτύου δε μπορεί να χρησιμοποιηθεί πλήρως, και το πρωτόκολλο μεταφοράς γίνεται το σημείο συμφόρησης του συνολικού συστήματος επικοινωνιών.

## 4.2 Δυνατότητα multicast

Σε πολλές εφαρμογές πολυμέσων απαιτείται η μετάδοση της ίδιας πληροφορίας σε πολλούς χρήστες ταυτόχρονα[1]. Σε αυτή της περίπτωση είναι χρήσιμο τα πρωτόκολλα πραγματικού χρόνου να υποστηρίζουν την multicast μετάδοση δεδομένων.

## 4.3 Χαρακτηριστικά QOS

Ροές δεδομένων πολυμέσων απαιτούν συνολικές QOS εγγυήσεις σχετικά με το εύρος ζώνης, την καθυστέρηση, και την διαταραχή καθυστέρησης. Για να ικανοποιήσει αυτές τις απαιτήσεις, ένα σύστημα μεταφοράς πρέπει να παρέχει ένα μηχανισμό στις εφαρμογές ώστε να μπορούν να καθορίζουν και να διαπραγματεύονται QOS απαιτήσεις.

Οι QOS απαιτήσεις που δίνονται στο πρωτόκολλο μεταφοράς περνούν στο πρωτόκολλο επιπέδου δικτύου. Το πρωτόκολλο επιπέδου δικτύου, το οποίο καλείται πρωτόκολλο δέσμευσης, διαδίδει αυτές τις απαιτήσεις και δεσμεύει τους απαραίτητους πόρους πάνω από μια σύνδεση δικτύου. Αυτή η σύνδεση συχνά είναι μια multicast σύνδεση σε εφαρμογές πολυμέσων.

Η παροχή QOS εγγυήσεων απαιτεί τη συνεργασία όλων των υποσυστημάτων ενός συστήματος μεταφοράς, περιλαμβάνοντας διαχείριση πόρων, έλεγχο της πρόσβασης στο δίκτυο, και διαχείριση ουρών σε διακόπτες (switches) δικτύου. Το λειτουργικό σύστημα θα πρέπει να υποστηρίζει εφαρμογές πολυμέσων.

## 5. Πρωτόκολλα δέσμευσης πόρων

---

Για την παροχή εγγυήσεων για QOS[5], πρέπει να χρησιμοποιηθούν τεχνικές διαχείρισης πόρων. Χωρίς διαχείριση πόρων σε τελικά συστήματα (αποστολείς και παραλήπτες των δεδομένων) δίκτυα και διακόπτες (switches), τα συστήματα πολυμέσων αδυνατούν να παρέχουν αξιόπιστη ποιότητα υπηρεσιών στους χρήστες. Μετάδοση δεδομένων πολυμέσων πάνω σε αδέσμευτους πόρους οδηγεί σε καθυστερημένα και χαμένα πακέτα λόγω της μη διαθεσιμότητας των απαιτούμενων πόρων. Συνεπώς η διαχείριση πόρων παίζει σημαντικό ρόλο στα συστήματα επικοινωνίας πολυμέσων. Για το λόγο αυτό ένα πολύ σημαντικό κομμάτι των συστημάτων επικοινωνίας πολυμέσων είναι το πρωτόκολλο δέσμευσης πόρων στο επίπεδο δικτύου. Αξιίζει να σημειωθεί ότι ένα πρωτόκολλο δέσμευσης πόρων δεν κάνει καμία δέσμευση απαιτούμενων πόρων το ίδιο, αλλά αποτελεί απλώς ένα μέσο που μεταφέρει πληροφορίες σχετικά με τις απαιτήσεις πόρων και διαπραγματεύεται τις τιμές του QOS που επιθυμούν οι χρήστες για τις end-to-end εφαρμογές τους. Τα πρωτόκολλα δέσμευσης βασίζονται σε συναρτήσεις διαχείρισης πόρων σε κάθε υποσύστημα για να επιτύχουν και να προγραμματίσουν προσβάσεις σε πόρους κατά τη διάρκεια της φάσης μετάδοσης.

Αν και η διαβάθμιση των μέσων, η οποία ρυθμίζει την ποσότητα των multimedia δεδομένων σύμφωνα με τους πόρους τους συστήματος που είναι κάθε στιγμή διαθέσιμοι, χρησιμοποιείται σε μερικά συστήματα, κάποια εγγύηση εξακολουθεί να απαιτείται από το σύστημα. Τα μέσα μπορούν να διαβαθμιστούν μόνο μέχρι κάποιο ορισμένο βαθμό και όχι κατά βούληση. Τα συστήματα επικοινωνίας πολυμέσων πρέπει να συνδυάζουν τεχνικές διαχείρισης πόρων και διαβάθμισης των μέσων για να πετύχουν την καλύτερη ποιότητα υπηρεσιών με μια συγκεκριμένη ποσότητα πόρων.

### 5.1 RSVP (Resource ReSerVation Protocol)

Το RSVP επιτρέπει στον παραλήπτη να ζητήσει μία ορισμένη end-to-end ποιότητα υπηρεσίας[6]ι εφαρμογές πραγματικού χρόνου χρησιμοποιούν το RSVP για να δεσμεύσουν τους απαραίτητους πόρους στους routers κατά μήκος του μονοπατιού μετάδοσης, έτσι ώστε να είναι διαθέσιμη η απαιτούμενη χωρητικότητα όταν λάβει χώρα η μετάδοση.

Το RSVP αποτελεί ένα πρωτόκολλο για multicasting και unicasting σηματοδότηση το οποίο σχεδιάστηκε για την εγκατάσταση και την συντήρηση σταθμών πληροφοριών σε κάθε router που βρίσκεται στο μονοπάτι μετάδοσης δεδομένων, κατά την μετάδοση δεδομένων. Ο χειρισμός καταστάσεων του RSVP ορίζεται στα RFC 2211 και RFC 2212 τα οποία έχουν οριστεί από την Integrated Services Work Group.

Για τη μετάδοση δεδομένων πολυμέσων ή δεδομένων υπερκειμένων πάνω από δίκτυο είναι αναγκαίο να ικανοποιούνται τρία βασικά χαρακτηριστικά:

1. Η μεταφορά των δεδομένων να γίνεται με όσο το δυνατό πιο γρήγορο τρόπο.
2. Να παρέχεται δυνατότητα multicasting (δηλαδή αποστολής δεδομένων σε πολλούς παραλήπτες με τη μεσολάβηση του δικτύου).
3. Να υπάρχει δυνατότητα για εξασφάλιση στην μεταφορά των δεδομένων με βάση τις απαιτήσεις που έχει ορίσει εκ των προτέρων ο χρήστης.

Τα δεδομένα υπερμέσων είναι μεγάλα σε όγκο και επομένως αποδοτικοί μηχανισμοί αποστολής τέτοιων δεδομένων πρέπει να παρέχονται. Το RSVP δείχνει περισσότερο ενδιαφέρον στη διατήρηση των παρεχόμενων πόρων και δεν μπορεί να επέμβει στη δρομολόγηση των δεδομένων που έχουν αποσταλεί.

### 5.1.1 Ανάπτυξη

Το RSVP ήταν το αποτέλεσμα μίας συλλογικής προσπάθειας από τους: Xerox Corp.'s Palo Alto Research Center (PARK), MIT, και Information Sciences Institute of University California (ISI).

### 5.12 Χαρακτηριστικά του RSVP

Το RSVP έχει τα παρακάτω χαρακτηριστικά:

Η ροή δεδομένων στο RSVP είναι μονής κατεύθυνσης. Το πρωτόκολλο διαχωρίζει τους αποστολείς από τους παραλήπτες. Παρόλο που σε πολλές περιπτώσεις ο αποστολέας μπορεί να είναι και παραλήπτης, το RSVP δεσμεύει πόρους μόνο προς τη μία κατεύθυνση.

Το RSVP υποστηρίζει και multicast και unicast και προσαρμόζεται στις συνεχείς αλλαγές ενός δυναμικού περιβάλλοντος. Το RSVP είναι receiver-oriented και μπορεί να χειριστεί διαφορετικές κατηγορίες παραληπτών. Ο κάθε παραλήπτης είναι υπεύθυνος για να διαλέξει το δικό του επίπεδο QOS. Ο αποστολέας διαχωρίζει την κίνηση σε ξεχωριστά flows, ένα για κάθε διαφορετικό επίπεδο QOS. Το RSVP έχει καλή συμβατότητα. Τρέχει πάνω από IPv4 και IPv6.

Επιτρέπει μεταφορά δεδομένων σε ετερογενείς παραλήπτες, δηλαδή παρέχεται η δυνατότητα σε μια σύνδεση multicast αποστολής διαφορετικής ποσότητας (amount of multimedia data) δεδομένων σε παραλήπτες που ανήκουν στην ίδια multicast σύνοδο.

Επιτρέπει η δυναμική σύνδεση και αποσύνδεση παραληπτών σε multicast σύνοδο.

Είναι αναγκαίο να υπάρχει ενημέρωση για τους διαθέσιμους πόρους πριν γίνουν αλλαγές στην δρομολόγηση.

Το RSVP είναι συμπληρωματικό του IP ελέγχοντας τον τρόπο με τον οποίο το IP στέλνει τα πακέτα του. Προορίζεται κυρίως για έλεγχο των δεδομένων που αποστέλλονται και όχι για μεταφορά δεδομένων.

Χρησιμοποιώντας το RSVP ένας αποστολέας δε γνωρίζει ποιοι παραλαμβάνουν τα δεδομένα που αποστέλλει.

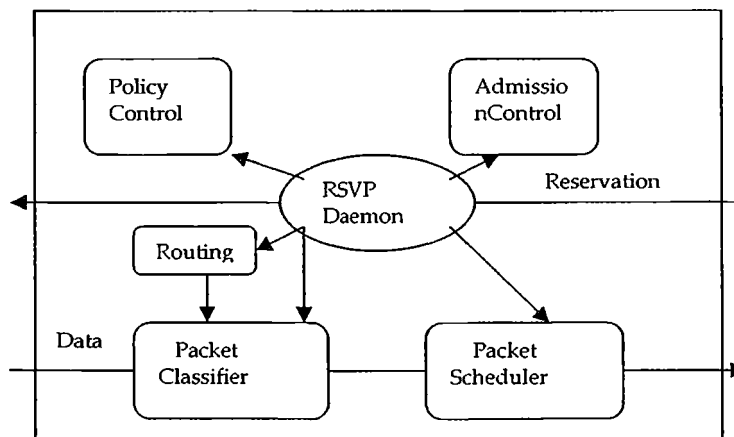
### 5.1.3 Πως δουλεύει το RSVP

Όταν μία εφαρμογή απαιτεί μία συγκεκριμένη ποιότητα υπηρεσίας, χρησιμοποιεί το πρωτόκολλο αυτό για να στείλει την απαίτηση σε όλους τους routers κατά μήκος του μονοπατιού μετάδοσης. Αν δεσμευτούν οι κατάλληλοι πόροι, το RSVP είναι υπεύθυνο για διατήρηση των πόρων αυτών.

Κάθε κόμβος που έχει την ικανότητα να δεσμεύει πόρους έχει αρκετές τοπικές διαδικασίες για το σκοπό αυτό. Το **policy control** καθορίζει αν ο χρήστης έχει το δικαίωμα να δεσμεύσει πόρους. Στο μέλλον, η διαδικασία αυτή θα περιλαμβάνει έλεγχο ταυτότητας, έλεγχο πρόσβασης και χρέωση. Το **admission control** ελέγχει τους πόρους και αποφασίζει για το αν ο κόμβος έχει αρκετούς πόρους για να υποστηρίξει την απαιτούμενη ποιότητα υπηρεσίας (QOS).

Το RSVP daemon πραγματοποιεί ελέγχους με βάση τις δύο αυτές διαδικασίες. Αν κάποιος από τους δύο ελέγχους αποτύχει, το RSVP πρόγραμμα επιστρέφει ένα error notification στην εφαρμογή που έκανε την αίτηση. Αν και οι δύο έλεγχοι είναι επιτυχημένοι, το RSVP daemon θέτει παραμέτρους στον packet classifier και στον packet scheduler έτσι ώστε να επιτευχθεί η ζητούμενη ποιότητα. Ο packet classifier καθορίζει την κλάση QOS για κάθε πακέτο και ο packet scheduler καθορίζει τη μετάδοση των πακέτων με στόχο την επίτευξη της ζητούμενης QOS για κάθε stream.

Το RSVP daemon επικοινωνεί επίσης με τη διαδικασία δρομολόγησης, για τον καθορισμό του μονοπατιού το οποίο θα ακολουθήσουν οι αιτήσεις δέσμευσης. Η δέσμευση γίνεται μέσω δύο τύπων μηνυμάτων του RSVP, τα PATH και τα RESV μηνύματα.



Σχήμα 5: Δέσμευση σε ένα κόμβο του μονοπατιού μετάδοσης των δεδομένων

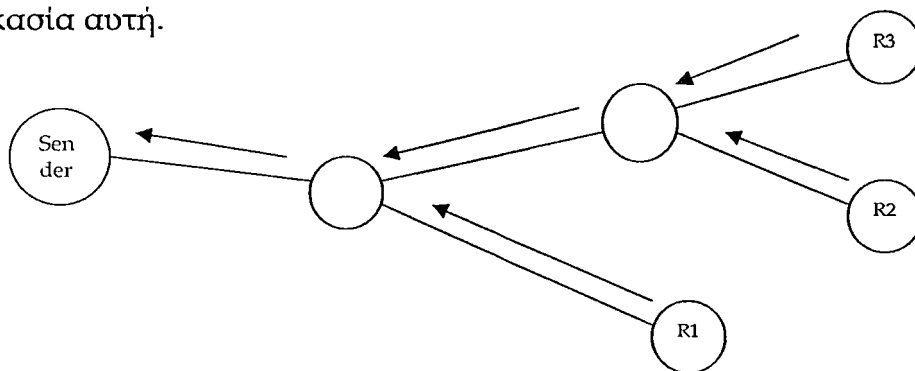


Τα PATH μηνύματα στέλνονται ανά τακτά χρονικά διαστήματα από τον αποστολέα στους παραλήπτες και περιλαμβάνουν το προφίλ των δεδομένων (data format, source address, source port) και άλλα χαρακτηριστικά για τη μεταφορά τους. Την πληροφορία αυτή χρησιμοποιούν οι παραλήπτες για να βρουν το αντίστροφο μονοπάτι προς τον αποστολέα και να προσδιορίσουν τους πόρους που πρέπει να δεσμευτούν.

Τα RESV μηνύματα δημιουργούνται από τους παραλήπτες και περιέχουν παραμέτρους για τη δέσμευση στις οποίες περιλαμβάνονται οι flow spec και filter spec. Η παράμετρος filter spec ορίζει ποια πακέτα πρέπει να χρησιμοποιηθούν από τον packet classifier. Η παράμετρος flow spec χρησιμοποιείται από τον packet scheduler. Τα RESV μηνύματα ακολουθούν το ακριβώς αντίθετο μονοπάτι των PATH μηνυμάτων.

Η δέσμευση που κάνει στους routers το RSVP καλείται soft states. Το RSVP daemon πρέπει ανά τακτά χρονικά διαστήματα να ανανεώνει τα μηνύματα έτσι ώστε να διατηρούνται οι δεσμευμένοι πόροι. Το γεγονός όμως αυτό καθιστά πιο εύκολες τις αλλαγές που είναι δυνατό να προκύψουν σε ένα δυναμικό περιβάλλον.

Οι αιτήσεις για δέσμευση αρχικοποιούνται από τους παραλήπτες. Δε χρειάζεται όλες οι αιτήσεις να ταξιδέψουν όλη τη διαδρομή μέχρι τον αποστολέα. Αντί γι' αυτό, τα streams που συναντώνται σε κάποιο κόμβο και κατευθύνονται στον ίδιο αποστολέα ενώνονται σε ένα stream, ενώνοντας και τις απαιτήσεις τους σε πόρους. Στο παρακάτω σχήμα φαίνεται πιο καλά η διαδικασία αυτή.



Σχήμα 6: Η ενοποίηση των αιτήσεων δέσμευσης

Το χαρακτηριστικό αυτό είναι και το πιο βασικό πλεονέκτημα του RSVP, που καλείται scalability. Ένας μεγάλος αριθμός χρηστών μπορεί να ενωθεί σε ένα multicast group χωρίς αυτό να αυξάνει σημαντικά το φόρτο του δικτύου.

Παρόλο που το RSVP βρίσκεται ένα επίπεδο πάνω από το IP, δεν είναι scheduling πρωτόκολλο. Στην πραγματικότητα, το RSVP βασίζεται στα routing πρωτόκολλα που βρίσκονται σε χαμηλότερα επίπεδα από αυτό.

Η αποστολή των παραμέτρων δέσμευσης (reservation parameters) είναι μια διαδικασία διαφορετική από αυτή του προσδιορισμού των παραμέτρων αυτών. Το RSVP είναι υπεύθυνο μόνο για την αποστολή των παραμέτρων αυτών.

Ο τρόπος με τον οποίο το RSVP προσθέτει ή απομακρύνει παραλήπτες από μια multicast σύννοδο είναι ο ίδιος τρόπος σύνδεσης με αυτόν που παρέχει το IP-multicast. Δηλαδή, κάποιος που θέλει να παρακολουθήσει τη σύννοδο μπορεί να ζητήσει να συνδεθεί. Στην περίπτωση αυτή, ο παραλήπτης προστίθεται σε μία λίστα από παραλήπτες που ήδη παρακολουθούν τη σύννοδο αυτή. Ο αποστολέας διαδίδει ένα κατάλληλο μήνυμα στο οποίο περιγράφονται οι απαιτήσεις σε πόρους. Μόλις ένας κόμβος πάρει ένα τέτοιο μήνυμα πρέπει να απαντήσει με ένα αντίστοιχο μήνυμα. Έτσι, επιτυγχάνεται μία σύνδεση. Σε περίπτωση που κάποιος θέλει να αποχωρίσει από τη σύννοδο μπορεί απλά να το κάνει στέλνοντας ένα κατάλληλο μήνυμα. Το μήνυμα μπορεί να σταλεί είτε από το παραλήπτη που θέλει να φύγει είτε από τον αποστολέα που θέλει να «διώξει» κάποιον παραλήπτη.

Ένα άλλο χαρακτηριστικό είναι η δυνατότητα για διαπραγμάτευση του παρεχόμενου επιπέδου εξυπηρέτησης από πολλαπλούς αποστολείς σε πολλαπλούς παραλήπτες. Το RSVP δίνει τη δυνατότητα σε ένα παραλήπτη να διατηρήσει μόνο ένα σύνολο από πόρους που μπορεί να χρησιμοποιηθεί από πολλούς αποστολείς. Ο παραλήπτης προσδιορίζει ποια πακέτα και από ποιους παραλήπτες θα “πάρει”. Με τον τρόπο αυτό, οι παραλήπτες μπορούν να “μεταπηδήσουν” από μία ροή δεδομένων σε μία άλλη.

Το RSVP δεν μπορεί να υποστηρίξει επίπεδο υπηρεσιών αφού δεν υπάρχει κάποια σύνδεση ανάμεσα στο τρόπο δρομολόγησης, τη δέσμευση πόρων και τη μεταφορά των δεδομένων. Το RSVP, λόγω της δυνατότητας για μεταπήδηση από μία ροή σε κάποια άλλη, είναι ιδανικό για εφαρμογές που κάνουν μετάδοση δεδομένων σε πολλούς χρήστες, του οποίου δε γνωρίζουν ούτε πόσοι είναι ούτε που είναι.

## 5.2 ST-II

Το ST-II είναι ένα πρωτόκολλο του Internet που σχεδιάστηκε το 1990 για να καταδίδει δεδομένα πραγματικού χρόνου. Είναι διάδοχος του ST πρωτοκόλλου το οποίο καθορίστηκε το 1979. Το ST-II υλοποιήθηκε για επικοινωνίες ήχου και κινούμενης εικόνας (video) σε διαφορετικά δίκτυα. Το ST-II παρέχει μια υπηρεσία προσανατολισμένη κατά σύνδεση (connection oriented) και υποστηρίζει multicast και διαπραγμάτευση QOS. Τα δεδομένα μπορούν να αποστέλλουν μόνο μετά την επιτυχή δημιουργία ενός stream (μια από-έναν-σε-πολλούς σύνδεση, ισοδύναμη με μια ροή).

### 5.2.1 Μορφή ST-II σύνδεσης

Η μορφή της σύνδεσης και η δέσμευση πόρων στο ST-II πρωτόκολλο υλοποιούνται ως απλά stream δεδομένων που σχηματίζουν ένα δέντρο multicast διανομής που έχει ως ρίζες της πηγές των δεδομένων και εκτείνεται μέχρι όλους του παραλήπτες.

Η δημιουργία ενός stream ξεκινάει όταν ένας κόμβος φορέας του ST, ο οποίος βρίσκεται στην πηγή του stream, δημιουργεί και στέλνει ένα Connect πακέτο το οποίο περιέχει τις προδιαγραφές του stream και το αρχικό σύνολο των συμμετεχόντων. Κάθε ενδιαμέσος φορέας του ST επεξεργάζεται το πακέτο Connect για να καθορίσει τα υποδίκτυα της επόμενης ή μετάβασης που απαιτούνται έτσι ώστε να φτάσει σε όλους τους παραλήπτες, εγκαθιστά την κατάσταση multicast προώθησης και δεσμεύει πόρους επιπέδου δικτύου κατά μήκος κάθε υποδικτύου. Η πληροφορία μπορεί να φτάσει σε πολλαπλούς παραλήπτες χρησιμοποιώντας την IP multicasting τεχνική. Αν η ποσότητα των πόρων που πραγματικά δεσμεύονται κατά μήκος κάθε υποδικτύου είναι μικρότερη από την ποσότητα που απαιτήθηκε, αυτό σημειώνεται στο Connect πακέτο ενημερώνοντας τις προδιαγραφές του stream. Όταν ένας παραλήπτης δέχεται μια ένδειξη, πρέπει να καθορίσει αν θέλει να συμμετάσχει στην ομάδα και να επιστρέψει είτε ένα μήνυμα Αποδοχής (Accept) είτε ένα μήνυμα Άρνησης (Refuse) στην πηγή του stream. Στην περίπτωση της αποδοχής ο παραλήπτης μπορεί να μειώσει ακόμα περισσότερο την απαίτηση για πόρους ενημερώνοντας τις προδιαγραφές του stream που επιστρέφει.

Η πηγή του stream πρέπει να περιμένει για μια απάντηση Αποδοχής ή Άρνησης από κάθε αρχικό παραλήπτη πριν ξεκινήσει την μετάδοση των δεδομένων. Το ST-II συμπεριφέρεται σε όλο το stream ένα μονοπάτι ομογενούς διανομής. Όταν η πηγή δέχεται μια απαίτηση Αποδοχής με μειωμένες προδιαγραφές ροής πρέπει είτε να προσαρμοστεί στο χαμηλό QOS για όλο το stream είτε να απορρίψει την συμμετοχή του συγκεκριμένου παραλήπτη στην ομάδα στέλνοντας του ένα μήνυμα Αποσύνδεσης. Όλοι οι συμμετέχοντες δέχονται δεδομένα με τις ίδιες προδιαγραφές stream αν και οι ικανότητες απεικόνισης μπορεί να διαφέρουν από παραλήπτη σε παραλήπτη. Οι παραλήπτες των stream μπορούν να προστεθούν ή να διαγραφούν μετά από την αρχική δημιουργία των stream. Κάθε προσθήκη ενός παραλήπτη απαιτεί μια αλληλεπίδραση με την πηγή του stream για να ξεκινήσει την αποστολή ενός μηνύματος σύνδεσης με την τρέχουσα προδιαγραφή του stream. Αυτή η αλληλεπίδραση δεν καθορίζεται από τις προδιαγραφές του πρωτοκόλλου αλλά εκτελείται χρησιμοποιώντας το IP. Όπως και στην αρχική προετοιμασία, η πηγή του stream πρέπει να εξετάσει τις προδιαγραφές του stream σε ένα μήνυμα αποδοχής που έχει επιστραφεί και είτε να δεχθεί τον νέο παραλήπτη είτε να τον απορρίψει αν οι πόροι που έχουν κατανεμηθεί είναι λιγότεροι από αυτούς που έχουν προς το παρόν δεσμευτεί για το stream.

Η διαγραφή των παραληπτών μπορεί να γίνει ασύγχρονα με την αποστολή ενός μηνύματος απόρριψης από έναν παραλήπτη ή με την αποστολή ενός μηνύματος αποσύνδεσης από την πηγή: το μήνυμα αποσύνδεσης μπορεί είτε να αναφέρει συγκεκριμένους παραλήπτες που πρέπει να διαγραφούν είτε να θέτει το flag σφαιρικής αποσύνδεσης ώστε να διακοπεί η μετάδοση του stream.

Στην περίπτωση φυσιολογική μετάδοσης, κατά την οποία δε σημειώνεται αποτυχία δρομολόγησης, τα stream που έχουν εγκαθιδρυθεί χρησιμοποιώντας το ST-II μπορούν να εξυπηρετηθούν έχοντας συγκεκριμένα χαρακτηριστικά QOS. Παρόλα αυτά, σημειώστε ότι το ST-II απλώς παρέχει το μηχανισμό για το πέρασμα και τη διαπραγμάτευση των QOS χαρακτηριστικών. Η πραγματική δέσμευση πόρων και ο πραγματικός χρονοπρογραμματισμός δεν είναι μέρος των προδιαγραφών του ST-II. Αυτά πρέπει να υλοποιούνται στον κάθε κόμβο του ST-II.

Ως ένα πρωτόκολλο προσανατολισμένο προς τη σύνδεση, το ST-II λαμβάνει τις αποφάσεις δρομολόγησης κατά τη διάρκεια της προετοιμασίας της σύνδεσης. Τα δεδομένα δεν προστατεύονται από αθροίσματα ελέγχου.

### 5.2.2. Χαρακτηριστικά του stream που μεταδίδει το ST-II

Το πρώτο βήμα για τον καθορισμό των QOS εγγυήσεων είναι ο καθορισμός των προδιαγραφών του stream. Οι προδιαγραφές καθορίζουν ως ένα μεγάλο βαθμό την τελική QOS και χρησιμοποίηση των πόρων του συστήματος που μπορούν να επιτευχθούν. Στο σχήμα 7 φαίνεται το πακέτο μέσο του οποίου καθορίζονται τα χαρακτηριστικά ενός stream το οποίο μεταδίδεται με το ST-II.

Για μια σύνδεση πραγματικού χρόνου, ενδιαφερόμαστε κυρίως για εγγυημένο εύρος ζώνης, άνω φράγματα καθυστέρησης και διαταραχής καθυστέρησης, και ρυθμό σφαλμάτων δεδομένων. Οι προδιαγραφές ενός ST-II stream καθορίζονται ως εξής: το ελάχιστο εύρος ζώνης καθορίζεται από το γινόμενο των `LimitOnPDUBytes` και `LimitOnPDURate`. Το άνω όριο της καθυστέρησης καθορίζεται από την παράμετρο `LimitOnDelay`. Δεν υπάρχει ρητή παράμετρος που να καθορίζει το άνω όριο της διαταραχής καθυστέρησης. Θεωρητικά το άνω όριο της διαταραχής ισούται με το άνω όριο της καθυστέρησης. Στην πράξη, η παράμετρος `AccDelayVariance` υποδεικνύει την αναμενόμενη διαταραχή καθυστέρησης. Τα ανώτερα στρώματα καθορίζουν το αν αυτή η τιμή είναι αποδεκτή. Οι εφαρμογές μπορούν να κατανέμουν το μέγεθος της ενδιάμεσης μνήμης εξομάλυνσης διαταραχής καθυστέρησης βασιζόμενες σε αυτή την τιμή. Ο μέγιστος ρυθμός σφαλμάτων bits καθορίζεται από την παράμετρο `ErrorRate` και ο ρυθμός απώλειας πακέτων υποδεικνύεται από την παράμετρο `Reliability`. Για μια VBR εφαρμογή που χρησιμοποιεί το ST-II, η πιο ενδιαφέρουσα λειτουργία είναι ο καθορισμός του ελάχιστου εύρους ζώνης που απαιτείται ώστε να υποστηρίζονται οι QOS απαιτήσεις της χωρίς να δεσμεύονται υπερβολικοί πόροι δικτύου.

bits: 0		16		31
PCode	PBytes	Version=3		0
Duty Factor	ErrorRace	Precedence		Reliability
Tradeoffs		Recovery Timecut		
LimitOnCost		LimitOnDelay		
LimitOnCost		LimitOnDelay		
LimitOnPDFUBytes		LimitOnPDURate		
MiniBytesXRate				
AccdMeanDelay				
AccdDelayVariance				
DesPDUBytes		DesPDURate		

**Σχήμα 7:** Δέσμευση πόρων με το ST-II

Από τη στιγμή που θα εγκαθιδρυθεί ένα stream, μονάδες δεδομένων του πρωτοκόλλου μπορούν να μεταφερθούν πάνω από αυτήν. Η επικεφαλίδα της ST PDU φαίνεται στο σχήμα 8. Σημειώστε ότι η μορφή της PDU είναι κοινή για όλες τις ST εκδόσεις και ότι ο αριθμός της έκδοσης υποδεικνύεται από ένα πεδίο έκδοσης.

Bits: 0		7		15		31
ST	VER	Pri	T	not used	TotalBytes	
HID			HeaderChecksum			
Timestamp{optional}						

**Σχήμα 8:** Επικεφαλίδα της ST-II PDU

Η ST-II PDU είναι αρκετά αποδοτική λόγω του header identifier (HID) που χρησιμοποιείται αντί για τον συνδυασμό των διευθύνσεων της πηγής και του προορισμού. Παρέχει οχτώ διαφορετικά επίπεδα προτεραιότητας. Η δυνατότητα χρησιμοποίησης διαφορετικών προτεραιοτήτων για διαφορετικά δεδομένα είναι σημαντική για επικοινωνίες πολυμέσων.

### 5.3 Σύγκριση RSVP και ST-II

Παρόλο που και τα δύο έχουν θα κοινό σκοπό την εξασφάλιση επικοινωνίας με QOS χαρακτηριστικά ανάμεσα σε πολλούς χρήστες, διαφέρουν σε πολλά σημεία[1]

Πρώτη μείζων διαφορά είναι στις προσεγγίσεις τους στην αρχή της δέσμευσης. Στο ST-II η δέσμευση ξεκινάει μεταφέροντας ο αποστολέας μια ροή προδιαγραφών στους αποδέκτες. Κατά τη διάρκεια της διάδοσης της ροής προδιαγραφών αυτής μερικοί παράμετροι είναι τροποποιημένοι στους διαθέσιμους πόρους στο path. Η τελική ροή προδιαγραφών που χρησιμοποιείται από τον αποστολέα για μεταφορά δεδομένων έχει τις ελάχιστες απαιτήσεις πόρων που είναι σύμφωνες με όλους τους δέκτες και το δίκτυο. Αφού η ροή έχει επαληθευθεί όλοι οι δέκτες λαμβάνουν δεδομένα με τη ίδια ροή προδιαγραφών αδιαφορώντας για τη δυναμικότητα υποστήριξης. Η σημασία αυτής της διαφοράς είναι ότι το RSVP είναι περισσότερο αποτελεσματικό στη χρήση πόρων δικτύου για τους ακόλουθους λόγους.

Αρχικά, στο RSVP, κάθε δέκτης λαμβάνει μόνο το σύνολο των δεδομένων που χρειάζεται έτσι που πόροι να μην εξαντλούνται. Η ροή στο ST-II είναι μια «έναν σε πολλούς» σύνδεση ξεκινώντας από ένα ειδικό αποστολέα. Αν περισσότεροι από ένας αποστολείς χρειάζεται να μεταβιβάσουν δεδομένα, κάθε αποστολέας πρέπει να χρησιμοποιήσει μια πολλαπλή σύνδεση. Έτσι κάθε παραλήπτης έχει πολλαπλές συνδέσεις, μια για κάθε αποστολέα. Αυτές οι συνδέσεις δεσμεύουν πόρους αλλά δεν μπορούν να είναι ενεργές όλη την ώρα, εξαντλώντας πόρους. Στο RSVP η ροή είναι μια «πολλοί σε πολλούς» σύνδεση. Αυτή η σύνδεση μπορεί να χρησιμοποιηθεί από διαφορετικούς αποστολέες εναλλακτικά ή σε συνδυασμό. Για εφαρμογές, όπως τηλεδιάσκεψη, που όλοι οι αποστολείς δεν είναι ενεργοί τη ίδια ώρα, το RSVP επιτρέπει πολύ αποτελεσματική χρήση των πόρων του δικτύου, ειδικά όταν το multicast σύνολο είναι μεγάλο.

Η δεύτερη διαφορά μεταξύ του ST-II και του RSVP είναι στις θέσεις ή το ρόλο στη στοίβα των πρωτοκόλλων. Το ST-II είναι ένα πλήρες δικτυακό πρωτόκολλο που περιέχει μεταχείριση δεδομένων όπως επίσης και λειτουργίες δέσμευσης πόρων. Αυτό αντικαθιστά το IP στο δικτυακό επίπεδο. Το RSVP, από την άλλη πλευρά, είναι ένα συνοδευτικό πρωτόκολλο του IP που ελέγχει τις διαδρομές και τη μεταφορά δεδομένων. Το πλεονέκτημα της προσέγγισης του RSVP είναι ότι το RSVP μπορεί να δουλέψει με διαφορετικές διαδρομές και πρωτόκολλα μεταφοράς δεδομένων. Αλλά το πρόβλημα είναι πώς να εξασφαλιστούν τα άλλα πρωτόκολλα, όπως το IP, ότι μπορούν να χρησιμοποιούν τους δεσμευμένους από το RSVP πόρους.

Η τρίτη διαφορά είναι στον τρόπο διαχείρισης πληροφορίας στην κατάσταση σύνδεσης. Το ST-II είναι ένα προσανατολισμένο στη σύνδεση. Όμοια με το ST-II, το RSVP αποθηκεύει σε κάθε κόμβο συμμετοχής πληροφορίας σχετικές με την υπάρχουσα ροή. Ωστόσο αυτή η πληροφορία καλείται soft-state που σημαίνει ότι η πληροφορία κατάστασης διατηρείται πάνω σε timeout βάση και χρειάζεται περιοδική ανανέωση.

Για ανακεφαλαίωση, το ST-II και το RSVP προσεγγίζουν εντελώς διαφορετικά τη δέσμευση πόρων. Το RSVP θα μπορούσε να χρησιμοποιεί πόρους δικτύου περισσότερο αποτελεσματικά, ειδικά όταν το multicast σύνολο είναι μεγάλο. Αυτά τα ζητήματα εξετάζονται ενεργά από την Internet Engineering Task Force (IETF) και άλλα ερευνητές.

## 6. Πρωτόκολλα Μεταφοράς

### 6.1 RTP/RTCP (Real-time Transport Protocol/Real-Time Control Protocol)

Τα πρωτόκολλα RTP/RTCP [7]ιουργήθηκαν για να χρησιμοποιηθούν σε εφαρμογές που απαιτούν μεταφορά ήχου, video και δεδομένων. Παρέχουν μια κοινή πλατφόρμα για την έκφραση πληροφοριών συγχρονισμού και συνόδων που απαιτούνται από εφαρμογές πραγματικού χρόνου. Το πρωτόκολλο RTCP αποτελεί το πρωτόκολλο ελέγχου του RTP. Το RTP είναι σχεδιασμένο να λειτουργεί σε συνεργασία με το πρωτόκολλο ελέγχου RTCP για να παίρνει πληροφορία για την ποιότητα της μετάδοσης και για αυτούς που συμμετέχουν στην επικοινωνία.

Το Real-time Transport Protocol (RTP) είναι ένα IP-based πρωτόκολλο για τη μεταφορά δεδομένων πραγματικού χρόνου όπως βίντεο και ήχος. Είναι σχεδιασμένο για multicast αλλά μπορεί να χρησιμοποιηθεί και για unicast. Μπορεί να χρησιμοποιηθεί για μονόδρομη επικοινωνία όπως video on demand, αλλά και για αμφίδρομη επικοινωνία όπως για διαδικτυακή τηλεφωνία.

Το RTP, είναι ένα πρωτόκολλο που προσφέρει end-to-end υπηρεσίες μεταφοράς για δεδομένα με χαρακτηριστικά πραγματικού χρόνου (real-time characteristics), όπως ήχος ή κινούμενη εικόνα (video) και άλλες εφαρμογές πάνω από δίκτυα εναλλαγής πακέτου, όπως το IP. Τέτοιες υπηρεσίες είναι ο καθορισμός και η αναγνώριση του τύπου των δεδομένων φόρτου (payload type), σειριακή αρίθμηση των πακέτων, χρονοσφράγιση πακέτων (timestamping) και έλεγχος των διαδικασιών μεταφοράς. Μια εφαρμογή θα χρησιμοποιεί το RTP πάνω από το TCP/IP ώστε να χρησιμοποιεί τις ευκολίες που παρέχει, ωστόσο μπορεί να χρησιμοποιηθεί πάνω από κάποιο άλλο κατάλληλο network ή transport protocol. Το RTP υποστηρίζει μεταφορά δεδομένων προς πολλαπλές κατευθύνσεις (multicast distribution) αν αυτό υποστηρίζεται από το δίκτυο.

Πρέπει να τονιστεί ότι το RTP δεν παρέχει κανένα μηχανισμό που εξασφαλίζει μεταφορά στα χρονικά όρια ούτε παρέχει εγγύηση για ποιότητα μετάδοσης (QOS). Αυτό είναι κάτι που αφορά τα πιο κάτω επίπεδα του δικτύου. Η αρίθμηση που παρέχεται στα πακέτα επιτρέπει στον παραλήπτη να διατάξει τα πακέτα στη σειρά που αυτά έφυγαν από τον αποστολέα.

Παρόλο που το κύριο πεδίο εφαρμογής για το οποίο είναι αρχικά σχεδιασμένο το RTP είναι η ικανοποίηση των αναγκών πολυμελούς τηλεδιάσκεψης πολυμέσων, εντούτοις δεν περιορίζεται στη συγκεκριμένη εφαρμογή. Εφαρμογές αποθήκευσης continues δεδομένων, interactive distributed simulation, active badge, εφαρμογές ελέγχου και μετρήσεων και άλλες εφαρμογές πραγματικού χρόνου μπορούν να χρησιμοποιήσουν το RTP ικανοποιητικά.



Το RTP μπορεί να χωριστεί σε δύο στενά συνδεδεμένα κομμάτια:

Το Real-time Transport Protocol (RTP), για μεταφορά δεδομένων με χαρακτηριστικά πραγματικού χρόνου και το Real-time Transport Control Protocol (RTCP), για έλεγχο της ποιότητας της υπηρεσίας και καταγραφή πληροφορίας σχετική με τα μέρη κάποιας ενεργού συνόδου. Το RTCP παρέχει λειτουργίες υποστήριξης για τηλεδιάσκεψη πραγματικού χρόνου για μεγάλες ομάδες στο Internet που περιλαμβάνουν αναγνώριση της πηγής και υποστήριξη για gateways (όπως audio & video bridges).

Το RTP παρουσιάζεται σαν ένα νέο "στυλ" πρωτοκόλλου, με την έννοια ότι μπορεί να παρέχει με εύκολο τρόπο την πληροφορία που απαιτεί μια εφαρμογή και συχνά χρησιμοποιείται σαν ένα τμήμα της εφαρμογής παρά σαν ένα ξεχωριστό επίπεδο. Επιπλέον μεταβολές ή προσθήκες στη μορφή του πρωτοκόλλου μπορούν εύκολα να γίνουν αλλάζοντας τη μορφή της επικεφαλίδας (header), σε αντίθεση με άλλα πρωτόκολλα όπου επιπρόσθετες θα πρέπει να εισαχθούν με τρόπο που να κάνει το πρωτόκολλο πιο γενικό ή να προστεθούν μηχανισμοί συντακτικού ελέγχου των αντίστοιχων πακέτων.

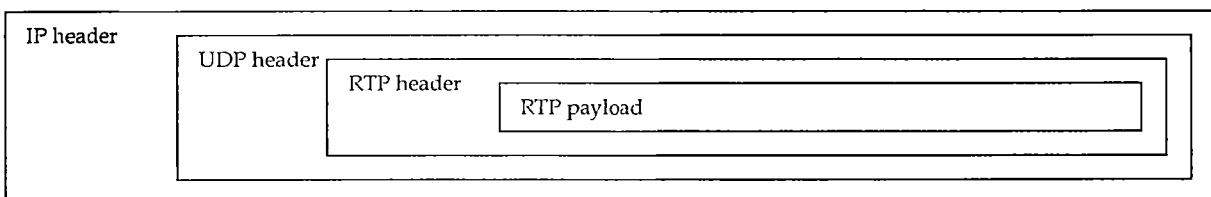
### 6.1.1 Πως λειτουργεί το RTP

Οι πολυμεσικές εφαρμογές χαρακτηρίζονται από αυστηρούς χρονικούς περιορισμούς στη μετάδοση των δεδομένων, κάτι που δε συμβαδίζει με τη λογική λειτουργίας του Internet[8]ο RTP παρέχει κάποιους μηχανισμούς που λαμβάνουν υπόψη τα θέματα αυτά. Τέτοιοι μηχανισμοί είναι το timestamping και το sequence numbering.

Το timestamping είναι η πιο σημαντική πληροφορία για τις εφαρμογές πραγματικού χρόνου. Ο αποστολέας βάζει σε κάθε πακέτο ένα timestamp, το οποίο χρησιμοποιεί ο παραλήπτης για να βρει τη χρονική στιγμή που πρέπει να «παιξουν» τα δεδομένα. Το timestamp χρησιμοποιείται επίσης για το συγχρονισμό διαφορετικών streams, όπως video και ήχου. Το RTP όμως δεν είναι υπεύθυνο για το συγχρονισμό αυτό, τον οποίο πραγματοποιούν οι εφαρμογές.

Το UDP δεν παραδίδει τα πακέτα με τη σειρά με την οποία στήθηκαν γι' αυτό τα πακέτα αριθμούνται, τη στιγμή που στέλνονται, έτσι να μπορεί ο παραλήπτης να τα βάλει στη σωστή σειρά. Οι αριθμοί χρησιμοποιούνται επίσης για να ανιχνεύονται απώλειες στη μετάδοση των πακέτων.

Ένα άλλο χαρακτηριστικό του RTP είναι το source identification, που επιτρέπει στον παραλήπτη να ξέρει από πού έρχονται τα δεδομένα. Οι παραπάνω μηχανισμοί πραγματοποιούνται μέσω του RTP header. Στο παρακάτω σχήμα φαίνεται ένα RTP πακέτο το οποίο έχει ενσωματωθεί σε ένα UDP/IP πακέτο.



Σχήμα 9: RTP δεδομένα σε ένα IP πακέτο

Το RTP τρέχει πάνω από το UDP. Το UDP και το UDP, είναι τα πρωτόκολλα που χρησιμοποιούνται πιο πολύ στο Internet. Το TCP παρέχει connection-oriented και αξιόπιστη ροή δεδομένων ανάμεσα σε δύο hosts, ενώ το UDP παρέχει connectionless αλλά όχι αξιόπιστες υπηρεσίες. Το UDP επελέγη σαν το πρωτόκολλο μεταφοράς του RTP για δύο λόγους:

Πρώτον, το RTP είναι κυρίως σχεδιασμένο multicast κάτι το οποίο δε συμβαδίζει με το connection-oriented TCP.

Δεύτερον, για δεδομένα πραγματικού χρόνου, η αξιοπιστία δεν είναι τόσο σημαντική όσο η έγκαιρη μετάδοση. Η αξιόπιστη μετάδοση η οποία επιτυγχάνεται μέσω της επαναμετάδοσης των χαμένων πακέτων, δεν είναι επιθυμητή διαδικασία, αφού μπορεί να προκαλέσει υπερφόρτωση του δικτύου, και προβλήματα στη συνεχή μετάδοση των δεδομένων.

Τα πακέτα του RTP και του RTCP μεταφέρονται μέσω UDP/IP. Παρόλα αυτά έχουν γίνει προσπάθειες για να συνεργάζονται και με άλλα πρωτόκολλα μεταφοράς.

### 6.1.2 RTP fixed header fields

Ο header του RTP έχει τη μορφή που γίνεται στο ακόλουθο σχήμα.

0	1	2	3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1			
V=2	P	X	CC
	M	PT	sequence number
Timestamp			
Synchronization source (SSRC) identifier			
Contributing source (CSRC) identifiers			

Σχήμα 10: RTP Packet Header

Ποιο αναλυτικά τα πεδία είναι τα ακόλουθα:

**Padding (P):** 1 bit. Αν είναι 1, τότε ο fixed header ακολουθείται από ακριβώς ένα header extension.

**CSRC count (CC):** 4 bits. Ο αριθμός των CSRC (contributing source) identifiers που ακολουθούν το fixed header. Ο αριθμός είναι μεγαλύτερος από ένα αν το πακέτο περιέχει δεδομένα από πολλές πηγές.

**Market (M):** 1 bit. Επιτρέπει σε διάφορα events όπως τα frame boundaries να μαρκάζονται στο packet stream.

**Payload type (PT):** 7 bits. Δηλώνει το είδος των δεδομένων, και παρέχει έτσι πληροφορίες για το πώς πρέπει να «παιχτούν» τα δεδομένα αυτά.

**Sequence number:** 16 bits. Αυξάνεται κατά 1 για κάθε πακέτο που στέλνεται και χρησιμοποιείται από τον παραλήπτη για να βρει τη σωστή σειρά των πακέτων και να ανιχνεύσει τυχόν απώλειες.

**Timestamp:** 32 bits. Χρησιμοποιείται για συγχρονισμό από τον παραλήπτη.

**SSRC:** 32 bits. Δείχνει αν τα δεδομένα προέρχονται από μία πηγή, ή αν προέρχονται από περισσότερες από μία, δείχνει σε ποιο σημείο ενώθηκαν τα δεδομένα.

**CSRC list:** 0 έως 15 items, 32 bits το καθένα. Δείχνει τις πηγές από τις οποίες προέρχονται τα δεδομένα. Ο αριθμός τους δίνεται από το πεδίο CC.

## 6.2 RTCP (Real-Time Control Protocol)

Το RTCP είναι ένα πρωτόκολλο ελέγχου που σχεδιάστηκε για να συνεργάζεται με το RTP. Υπάρχουν οι παρακάτω 5 τύποι πακέτων για το πρωτόκολλο RTCP:

**RR:** receiver report. Περιέχουν πληροφορία για την ποιότητα των δεδομένων στα σημεία της παραλαβής τους, καθώς επίσης και στατιστικά στοιχεία.

**SR:** sender report. Δημιουργούνται από τους αποστολείς και περιέχουν πληροφορία τα δεδομένα που στέλνονται.

**SDES:** source description items. Περιέχουν πληροφορία για τις πηγές (sources) των δεδομένων.

**BYE:** δηλώνει τέλος συμμετοχής.

**APP:** application specific functions. Χρησιμοποιείται για πειραματικούς λόγους.

Μέσω των παραπάνω πακέτων ελέγχου, το RTCP παρέχει τις παρακάτω υπηρεσίες:

**QOS monitoring και congestion control.** Είναι η βασική συνάρτηση του RTCP. Το RTCP παρέχει πληροφορία (feedback) στις εφαρμογές για την ποιότητα των δεδομένων.

**Source identification.** Στα πακέτα δεδομένων του RTP, οι πηγές (sources) αναγνωρίζονται μέσω 32 bit identifiers οι οποίοι παράγονται τυχαία. Οι identifiers αυτοί δεν είναι βολικοί για τους ανθρώπους. Τα RTCP SDES (source description) πακέτα περιέχουν textual πληροφορία που καλείται canonical names, στη θέση των identifiers.

**Inter-media synchronization.** Τα RTCP sender reports περιέχουν ένα indication για το πραγματικό χρόνο και το αντίστοιχο RTP timestamp. Η πληροφορία αυτή μπορεί να χρησιμοποιηθεί για εσωτερικό συγχρονισμό (π.χ. εικόνα και ήχος σε ένα video stream).

**Control information scaling.** Τα RTCP πακέτα στέλνονται περιοδικά ανάμεσα σε αυτούς που συμμετέχουν στην επικοινωνία. Καθώς ο αριθμός των συμμετεχόντων αυξάνεται, γίνεται απαραίτητη η αποκατάσταση κάποιας ισορροπίας ανάμεσα στην πληροφορία ελέγχου που ανταλλάσσεται και στο φόρτο του δικτύου. Το RTCP στην περίπτωση αυτή μειώνει την πληροφορία ελέγχου μέχρι και 5%, έτσι ώστε να αποφευχθεί η υπερφόρτωση του δικτύου.

### 6.3 Χαρακτηριστικά του RTP/RTCP

Το RTP υπηρεσίες για την παράδοση δεδομένων με χαρακτηριστικά πραγματικού χρόνου, όπως video και ήχος, αλλά δεν παρέχει κανένα μηχανισμό που να εγγυάται την έγκαιρη μεταφορά. Χρειάζεται υποστήριξη από άλλα επίπεδα που έχουν τον έλεγχο των πόρων. Για τη δέσμευση των απαραίτητων πόρων, το RTP βασίζεται στο RSVP.

Το RTP δεν υποθέτει τίποτα για το δίκτυο, εκτός από το γεγονός ότι το δίκτυο παρέχει framing. Το RTP τρέχει πάνω από το UDP, αλλά έχουν γίνει προσπάθειες ώστε να γίνει συμβατό και με άλλα πρωτόκολλα.

Δεν παρέχει κανένος είδους αξιοπιστία. Υποστηρίζει timestamps, sequence numbers και flow/congestion control, αλλά η υλοποίηση τους αφήνεται αποκλειστικά για την εφαρμογή.

Το RTP δε μπορεί να θεωρηθεί ολοκληρωμένο, αλλά είναι ανοικτό στην εισαγωγή νέων formats και νέου λογισμικού πολυμέσων.

Το RTP/RTCP παρέχει functionality και μηχανισμούς ελέγχου για τη μεταφορά δεδομένων πραγματικού χρόνου. Δεν είναι όπως υπεύθυνο για τις δραστηριότητες υψηλότερο επιπέδου, όπως assembly και συγχρονισμός, οι οποίες γίνονται στο application level.

Το flow και congestion control information στο RTP παρέχεται από τα RTCP sender και receiver reports.

## 7. Πρωτόκολλο streaming

---

### 7.1 Το πρωτόκολλο RTSP (Real Time Streaming Protocol)

#### 7.1.1 Γενική περιγραφή

Το RTSP είναι ένα πρωτόκολλο επιπέδου εφαρμογής για τη διαχείριση της διανομής δεδομένων που έχουν streaming ιδιότητες. Στις δυνατότητές συγκαταλέγεται το γεγονός ότι παρέχει ένα επεκτάσιμο πλαίσιο ώστε να ενεργοποιεί την ελεγχόμενη, κατά παραγγελία (on demand) διανομή δεδομένων πραγματικού χρόνου όπως είναι το βίντεο και ο ήχος. Οι πηγές των δεδομένων μπορούν να περιέχουν τόσο τροφοδοσία ζωντανών δεδομένων όσο και αποθηκευμένων. Σκοπός του πρωτοκόλλου είναι:

- . να διαχειρίζεται πολλαπλές συνεδρίες διανομής δεδομένων
- . να παρέχει κάποιο μέσο για την επιλογή των καναλιών διανομής όπως είναι το UDP [7]
- . να κάνει multicast των UDP και TCP
- . και να παρέχει το μέσο για την επιλογή του μηχανισμού διανομής που βασίζεται πάνω στο RTP.

#### 7.1.2 Αναλυτικά οι αιτίες ύπαρξης του πρωτοκόλλου RTSP

Το RTSP εγκαθιστά και διαχειρίζεται είτε ένα είτε πολλαπλά χρονικά συγχρονισμένα streams συνεχών μέσων όπως είναι ο ήχος και το βίντεο. Τυπικά δεν διανέμει τα συνεχή ρεύματα δεδομένων αυτά καθαυτά, έστω κι αν είναι δυνατή η διαστρωμάτωση των συνεχών ρευμάτων μέσων με ρεύματα ελέγχου. Με άλλα λόγια το RTSP συμπεριφέρεται ως δικτυακό «τηλεκοντρόλ» για τους εξυπηρετητές πολυμέσων (multimedia servers).

Δεν υπάρχει καμία ένδειξη για κάποια RTSP σύνδεση, αντιθέτως ο εξυπηρετητής διατηρεί κάποια συνεδρία μαρκαρισμένη μέσω ενός δείκτη. Σε καμία περίπτωση δεν είναι κάποια RTSP συνεδρία στενά δεμένη με κάποια σύνδεση στο επίπεδο μεταφοράς π.χ. μια TCP σύνδεση. Κατά τη διάρκεια μιας RTSP πελάτης να ξεκινήσει ή να διακόψει πολλές αξιόπιστες συνδέσεις μεταφοράς με τον εξυπηρετητή με σκοπό να διανέμει τις RTSP αιτήσεις. Εναλλακτικά μπορεί να χρησιμοποιήσει ένα πρωτόκολλο μεταφοράς χωρίς σύνδεση (connectionless) όπως είναι το UDP. Τα διαχειριζόμενα από το RTSP, ρεύματα ροής μπορούν να χρησιμοποιήσουν το RTP, αλλά η λειτουργία του RTSP δεν εξαρτάται από τον μηχανισμό μεταφοράς που χρησιμοποιείται για την μεταφορά των συνεχών μέσων.

Το πρωτόκολλο έχει σκοπίμως παρόμοια σύνταξη και λειτουργία με το HTTP 1.1 έτσι ώστε οι μηχανισμοί επέκτασης του HTTP να μπορούν να προστεθούν στις περισσότερες περιπτώσεις και στο RTSP. Ωστόσο το RTSP διαφέρει από το HTTP σε έναν αριθμό από σημαντικά θέματα:

- Το RTSP εισάγει νέες μεθόδους και έχει διαφορετικούς δείκτες πρωτοκόλλου.
- Σε αντιδιαστολή με την άνευ κατάστασης φύση του HTTP σε όλες σχεδόν τις περιπτώσεις, πρέπει ένας RTSP εξυπηρετητής να διατηρεί την κατάσταση του.
- Στο RTSP μπορούν να αποστέλλουν αιτήσεις και ο πελάτης αλλά και ο εξυπηρετητής.
- Τα εκτός ζώνης (ή ομάδας) δεδομένα μεταφέρονται με άλλο πρωτόκολλο. (ωστόσο υπάρχει μια εξαίρεση σ' αυτό). Το RTSP έχει καθοριστεί να χρησιμοποιεί το ISO 10646 (UTF-8) αντί του ISO 8859-1 που είναι σύμφωνο με τις προσπάθειες για διεθνοποίηση του τωρινού HTML.
- Η Request-URI περιέχει πάντοτε την απόλυτη URI. Το HTTP/1.1 λόγω της συμβατότητας προς τα πίσω (backward compatibility) με ένα ιστορικό ατόπημα, μεταφέρει όνο το απόλυτο μονοπάτι στην αίτηση και τοποθετεί το όνομα του ξενιστή (host)<sup>7</sup> σε ένα ξεχωριστό πεδίο της επικεφαλίδας. Αυτό κάνει ευκολότερο το 'virtual hosting', όπου ένας ξενιστής με μοναδική IP διεύθυνση φιλοξενεί πολλές δενδρικές δομές αρχείων (document trees).

### 7.1.3 Λειτουργίες του πρωτοκόλλου

Το RTSP υποστηρίζει τις εξής λειτουργίες:

- Ανάκτηση του μέσου από έναν εξυπηρετητή μέσων (media server): ο πελάτης μπορεί να ζητήσει μια περιγραφική παρουσίαση των πληροφοριών μέσω του HTTP ή κάποιας άλλης μεθόδου. Αν η παρουσίαση έχει γίνει multicast, τότε η περιγραφή της περιέχει τις εν δυνάμει multicast διευθύνσεις και θύρες που μπορούν να χρησιμοποιηθούν από τα συνεχή δεδομένα. Αν η παρουσίαση είναι προς αποστολή μόνο στον πελάτη μέσω unicast, τότε ο τελευταίος παρέχει τον προορισμό για λόγους ασφαλείας.
- Πρόσκληση ενός εξυπηρετητή μέσων σε μια συνδιάσκεψη ένας εξυπηρετητής μέσων μπορεί να προσκληθεί για να πάρει μέρος σε μια συνδιάσκεψη, είτε για να παράγει κάποιο μέσο (π.χ. βίντεο ή ήχος) στην παρουσίαση, είτε για να καταγράψει ολόκληρα ή μέρος των μιας παρουσίασης. Αυτή η κατάσταση είναι χρήσιμη σε εφαρμογές καταμεμημένης διδασκαλίας. Πολλά μέλη μιας συνδιάσκεψης μπορούν να κατευθυνθούν αλλού «απλά πατώντας τα πλήκτρα του τηλεχειριστηρίου».
- Πρόσθεση μέσων σε μια ήδη δημιουργημένη παρουσίαση: είναι χρήσιμο, ειδικότερα για ζωντανές παρουσιάσεις, να μπορεί ένας εξυπηρετητής να πει στον πελάτη σχετικά με το αν έχουν γίνει διαθέσιμα πρόσθετα μέσα.

- Οι RTSP αιτήσεις μπορούν να διαχειριστούν από πληρεξούσιους (proxies), τούνελ (tunnels) και caches όπως γίνεται και στην περίπτωση του HTTP/1.1

#### 7.14 Ιδιότητες του πρωτοκόλλου

Το RTSP έχει τις εξής ιδιότητες:

- Είναι επεκτάσιμο. Μπορούν πολύ εύκολα να προστεθούν στο RTSP νέες μέθοδοι και παράμετροι.[7]
- Είναι εύκολο να μεταγλωττιστεί σε μια ενδιάμεση μορφή και αυτή η μορφή να μεταφράζεται στη συνέχεια από τους κατάλληλους μεταφραστές (parsers). Το RTSP μπορεί να γίνει parsed από τους συνήθεις HTTP ή MIME parsers.
- Είναι ασφαλές. Το RTSP επαναχρησιμοποιεί τους μηχανισμούς ασφαλείας του web. Όλοι οι HTTP μηχανισμοί πιστοποίησης όπως η βασική ή συνοπτική (digest) πιστοποίηση (authentication) είναι εφαρμόσιμοι απευθείας.
- Είναι ανεξάρτητο του επιπέδου μεταφοράς. Το RTSP μπορεί να χρησιμοποιήσει είτε unreliable datagram protocol (UDP) είτε το reliable datagram protocol (RDP) δεν είναι πολύ διαδεδομένο), είτε ένα αξιόπιστο stream πρωτόκολλο όπως το TCP.
- Είναι multi-server capable. Αυτό σημαίνει ότι κάθε ρεύμα μέσων που μετέχει σε μια παρουσίαση μπορεί να παραμένει σε διαφορετικό εξυπηρετητή. Ο πελάτης εγκαθιστά αυτόματα πολλά sessions διαχείρισης, με τους διάφορους εξυπηρετητές μέσων, που εκτελούνται παράλληλα. Ο συγχρονισμός των μέσων εκτελείται στο επίπεδο μεταφοράς.
- Διαχειρίζεται τις συσκευές εγγραφής. Το πρωτόκολλο μπορεί να διαχειριστεί τόσο τις συσκευές αναπαραγωγής όσο και εγγραφής καθώς και συσκευές να εναλλάσσονται μεταξύ των δύο ρόλων.
- Διαχωρισμός μεταξύ της διαχείρισης του ρεύματος και της έναρξης της διάσκεψης. Η διαχείριση του ρεύματος έχει χωριστεί από την πρόσκληση κάποιου εξυπηρετητή μέσων σε μια συνδιάσκεψη. Η μόνη απαίτηση που υπάρχει είναι το πρωτόκολλο έναρξης είτε να παρέχει είτε να μπορεί να χρησιμοποιηθεί για να παράγει μοναδικούς δείκτες διάσκεψης. Συγκεκριμένα μπορεί να χρησιμοποιηθεί το SIP ή το H.323 (International Telecommunication Union, 1996) για να προσκληθεί ένας εξυπηρετητής σε κάποια συνδιάσκεψη.
- Είναι κατάλληλο για επαγγελματικές εφαρμογές. Το RTSP υποστηρίζει ακρίβεια σε επίπεδο πλαισίου (frame) μέσω των SMPTE χρονοσφραγίδων έτσι ώστε να επιτρέπει την απομακρυσμένη ψηφιακή επεξεργασία.
- Χρησιμοποιεί ουδέτερη περιγραφή παρουσίασης. Το πρωτόκολλο δεν επιβάλλει κάποια συγκεκριμένη μορφή περιγραφής της παρουσίασης ή κάποια ειδική metafile (συνοδευτικό αρχείο συμπληρωματικών

- πληροφοριών) μορφοποίηση (format) και μπορεί να μεταφέρει τον τύπο της μορφοποίησης που πρόκειται να χρησιμοποιηθεί τόσο η περιγραφή πρέπει να περιέχει τουλάχιστον μια RTSP URL.
- Είναι proxy και firewall friendly. Το πρωτόκολλο πρέπει να βρίσκεται σε ετοιμότητα ώστε να διαχειριστεί από του φράκτες (firewalls) τόσο επιπέδου εφαρμογής όσο και επιπέδου μεταφοράς. Ένας φράκτης πρέπει να μπορεί να κατανοήσει τη SETUP διαδικασία εστί ώστε να ανοίξει μια «τρύπα» για το UDP ρεύμα μέσων.
- Είναι HTTP friendly. Το RTSP όντας ευαίσθητο, επαναχρησιμοποιεί τις αρχές που διέπουν το HTTP έτσι ώστε να μπορεί να ξαναχρησιμοποιηθεί η υπάρχουσα υποδομή. Αυτή η υποδομή περιλαμβάνει PICS (Platform for Internet Content Selection) για την αντιστοίχιση ετικετών με το περιεχόμενο. Ωστόσο το RTSP δεν προσθέτει απλά μεθόδους στο HTTP μιας και η διαχείριση συνεχών μέσων απαιτεί στις περισσότερες περιπτώσεις να βρίσκεται στην πλευρά του εξυπηρετητή.
- Κάνει κατάλληλη διαχείριση του εξυπηρετητή. Αν ένας πελάτης μπορεί να ξεκινήσει ένα stream τότε πρέπει να έχει επίσης και τη δυνατότητα να το σταματήσει. Οι εξυπηρετητές δεν πρέπει να ξεκινούν ρεύματα συνεχούς ροής δεδομένων προς του πελάτες κατά τέτοιο τρόπο που να μην δίνεται στους πελάτες η δυνατότητα να τα σταματήσουν.
- Διαπραγματεύσεις μεταφοράς. Ο πελάτης μπορεί να διαπραγματευτεί κάποια μέθοδο προτού την χρειαστεί πραγματικά για να επεξεργαστεί κάποιο συνεχές ρεύμα μέσων.
- Δυνατότητα διαπραγμάτευσης. Πρέπει να υπάρχουν κάποιοι σαφείς μηχανισμοί από πλευράς του πελάτη που να καθορίζουν ποιες μέθοδοι δεν θα υλοποιηθούν σε περίπτωση που τα βασικά χαρακτηριστικά είναι απενεργοποιημένα. Κάτι τέτοιο επιτρέπει στους πελάτες να εμφανίζουν το κατάλληλο user interface. Αν για παράδειγμα δεν επιτρέπεται να αναζήτηση μέσα στην παρουσίαση τότε το GUI (Graphical User Interface) πρέπει να έχει τη δυνατότητα να απενεργοποιήσει την μετακίνηση του ολισθαίνοντα δείκτη θέσης της παρουσίασης.

#### 7.1.4 Η γενική λειτουργία του πρωτοκόλλου

Κάθε παρουσίαση και κάθε ρεύμα μέσων μπορεί να προσδιοριστεί από μια RTSP URL. Η συνολική παρουσίαση και οι ιδιότητες των μέσων από τα οποία αποτελείται είναι καθορισμένα από το αρχείο περιγραφής. Αυτό το αρχείο μπορεί να αποκτηθεί από τον πελάτη χρησιμοποιώντας το HTTP ή άλλα μέσα όπως το ηλεκτρονικό ταχυδρομείο (email) και μπορεί να μην αποθηκευθεί απαραίτητα στον εξυπηρετητή μέσων. Για να ικανοποιούνται αυτές οι προδιαγραφές θεωρείται ότι η περιγραφή παρουσίασης (presentation description) περιγράφει μία ή περισσότερες παρουσιάσεις, καθεμία των οποίων διατηρεί έναν κοινό χρονικό άξονα.



Θεωρούμε για λόγους απλότητας και χωρίς να επηρεαστεί η γενικότητα ότι η περιγραφή παρουσίασης περιέχει ακριβώς μία τέτοια παρουσίαση.

Μια παρουσίαση μπορεί να περιέχει πολλά ρεύματα μέσω. Το αρχείο περιγραφής της περιέχει τις περιγραφές των ρευμάτων που την συνθέτουν περιλαμβάνοντας και τις κωδικοποιήσεις τους, την γλώσσα καθώς και άλλες παραμέτρους που εξουσιοδοτούν τον πελάτη να διαλέξει τον καταλληλότερο συνδυασμό των μέσω. Μέσα σε αυτή την περιγραφή παρουσίασης κάθε ρεύματος μέσω, που είναι ξεχωριστά διαχειριζόμενο από εξυπηρετητή μέσω που διαχειρίζεται το συγκεκριμένο ρεύμα μέσω και ονομάζει το ρεύμα που είναι αποθηκευμένο σ' αυτόν. Διάφορα ρεύματα μέσω μπορούν να τοποθετηθούν σε διαφορετικούς εξυπηρετητές. Για παράδειγμα τα ρεύματα ήχου και βίντεο μπορούν να χωριστούν σε διάφορους εξυπηρετητές έτσι ώστε να φορτώνονται κατανεμημένα. Εκτός αυτών η περιγραφή απαριθμεί της μεθόδους μεταφοράς που έχει τη δυνατότητα να υποστηρίξει ο εξυπηρετητής. Είναι αναγκαίο να καθοριστούν, εκτός των παραμέτρων των μέσω, και η δικτυακή διεύθυνση προορισμού καθώς και οι θύρες που πρόκειται να χρησιμοποιηθούν. Μπορούμε να διακρίνουμε διάφορους τρόπους λειτουργίας:

- Unicast: Τα μέσα μεταδίδονται στην πηγή της RTSP αίτησης με τον αριθμό θύρας που επιλέχτηκε από τον πελάτη. Εναλλακτικά, τα μέσα μεταδίδονται με το ίδιο αξιόπιστο ρεύμα όπως και το RTSP.
- Multicast, με τον εξυπηρετητή να επιλέγει την διεύθυνση: Σ' αυτή την περίπτωση την multicast διεύθυνση και τη θύρα την επιλέγει ο εξυπηρετητής μέσω. Αυτή είναι η τυπική περίπτωση ζωντανής ή σχεδόν media-on-demand μετάδοσης.
- Multicast, όπου ο πελάτης επιλέγει την διεύθυνση: Αν ο εξυπηρετητής πρόκειται να συμμετάσχει σε μια υπάρχουσα multicast συνδιάσκεψη, η διεύθυνση, η θύρα και το κλειδί κρυπτογράφησης δίνονται από την περιγραφή της συνδιάσκεψης.

### 7.1.5 Καταστάσεις του RTSP

Το RTSP διαχειρίζεται ένα stream που μπορεί να έχει σταλεί μέσω κάποιου άλλου πρωτοκόλλου ανεξαρτήτου από το κανάλι διαχείρισης. Η RTSP διαχείριση, για παράδειγμα, μπορεί να προκύψει σε μια TCP σύνδεση τη στιγμή που τα δεδομένα ρέουν μέσω UDP. Για αυτό η διανομή δεδομένων συνεχίζεται ακόμα κι αν δεν παραλαμβάνονται RTSP αιτήσεις από τον εξυπηρετητή. Επίσης κατά τη διάρκεια ύπαρξής του, ένα ρεύμα μέσω μπορεί να διαχειρίζεται από RTSP αιτήσεις που εκπέμπονται αλυσιδωτά από διαφορετικές TCP συνδέσεις. Για αυτό τον λόγο χρειάζεται να διατηρεί ο εξυπηρετητής την κατάσταση της συνεδρίας έτσι ώστε να είναι σε θέση να συσχετίζει τις RTSP αιτήσεις με το stream.[8]

Πολλές μέθοδοι στο RTSP δεν συνεισφέρουν στην κατάσταση.

Στη συνέχεια παρατίθεται οι σημαντικότερες μέθοδοι για τον καθορισμό της δέσμευσης και χρήσης των πόρων του ρεύματος σε έναν εξυπηρετητή.

- **SETUP:** Προκαλεί τον εξυπηρετητή να δεσμεύσει πόρους για ένα stream και να ξεκινήσει μία RTSP συνεδρία.
- **PLAY και RECORD:** Ξεκινά τη μετάδοση δεδομένων ενός stream που δεσμεύτηκε μέσω της κατάστασης SETUP. Η αίτηση PLAY χρησιμοποιείται για την αναπαραγωγή του μέσου ενώ η RECORD για την εγγραφή και αποθήκευσή του.
- **PAUSE:** Σταματά προσωρινά κάποιο stream χωρίς να απελευθερώσει τους πόρους που έχει δεσμεύσει ο εξυπηρετητής για αυτό.
- **TEARDOWN:** Ελευθερώνει του πόρους που αντιστοιχούν στο stream.
- Η RTSP συνεδρία σταματά να υπάρχει στον εξυπηρετητή.

Οι RTSP μέθοδοι που συνεισφέρουν στην κατάσταση χρησιμοποιούν το session πεδίο της επικεφαλίδας για την ανίχνευση της RTSP συνεδρίας του οποίου η κατάσταση τελεί υπό διαχείριση. Ο εξυπηρετητής παράγει δείκτες συνεδρίας (session identifiers) σε απόκριση των SETUP αιτήσεων.

### 7.1.6 Σχέση με άλλα πρωτόκολλα

Στο RTSP υπάρχει μια αλληλοεπικάλυψη ως αναφορά στη λειτουργία του με το HTTP. Μπορεί επίσης να αλληλεπιδρά με το HTTP στο ότι η αρχική σύνδεση με streaming περιεχόμενο πρόκειται να πραγματοποιηθεί μέσω μιας σελίδας του web. Οι προδιαγραφές του πρωτοκόλλου στοχεύουν να επιτρέπουν διάφορα σημεία συναλλαγών μεταξύ ενός εξυπηρετητή web αναφέρουμε ότι η περιγραφή παρουσίασης μπορεί να ανακτηθεί είτε χρησιμοποιώντας RTSP είτε HTTP κάτι που περιορίζει τις παλινδρομήσεις στις περιπτώσεις που βασίζονται σε πλοηγητή ιστοσελίδων (web browser) και ισχύει ακόμα και στην περίπτωση που έχουμε μεμονωμένους (standalone) RTSP εξυπηρετητές και πελάτες που δεν στηρίζονται καθόλου στο HTTP.

Ωστόσο το RTSP διαφέρει ριζικά από το HTTP στο ότι η διανομή δεδομένων πραγματοποιείται με διαφορετικό πρωτόκολλο. Το HTTP είναι ένα ασύμμετρο πρωτόκολλο όπου ο πελάτης δημιουργεί αιτήσεις και ο εξυπηρετητής ανταποκρίνεται σ' αυτές. Στο RTSP μπορούν να δημιουργήσουν αιτήσεις τόσο ο πελάτης μέσω των μέσων όσο και ο εξυπηρετητής μέσω των μέσων. Επίσης οι RTSP αιτήσεις δεν στερούνται κατάστασης (stateless). Έτσι έχουν την δυνατότητα να συνεχίσουν να καθορίζουν παραμέτρους και να διαχειρίζονται κάποιο ρεύμα μέσω των μέσων για αρκετό χρονικό διάστημα μετά την βεβαίωση της λήψης της αίτησης. [7]

Η επαναχρησιμοποίηση της λειτουργικότητας του HTTP έχει πλεονεκτήματα σε τουλάχιστον δύο περιπτώσεις κι αυτές είναι η ασφάλεια και οι proxies. Υπάρχουν σχεδόν οι ίδιες απαιτήσεις και έτσι η δυνατότητα να χρησιμοποιηθεί ο τρόπος που λειτουργεί το HTTP στις caches (ενδιάμεσοι αποθηκευτικοί χώροι), στους proxies και στην πιστοποίηση είναι πολύτιμη. Παρόλο που τα περισσότερα μέσα πραγματικού χρόνου (real-time media) θα χρησιμοποιήσουν το RTP ως πρωτόκολλο μεταφοράς, το RTSP δεν συνδέεται στενά με αυτό. Το RTSP θεωρεί ότι υπάρχει μια μορφοποίηση περιγραφής παρουσίασης (presentation description format) που μπορεί να εκφράσει τόσο στατικές όσο και προσωρινές ιδιότητες μιας παρουσίασης που περιέχει διάφορα ρεύματα μέσων.

### 7.1.7 Συνδέσεις

Οι RTSP αιτήσεις μπορούν να μεταδοθούν με διάφορους τρόπους:

- Να χρησιμοποιούνται για διάφορες συναλλαγές αίτησης-απόκρισης συνεχείς συνδέσεις μεταφοράς.
- Να χρησιμοποιείται μια σύνδεση ανά συναλλαγή.
- Να χρησιμοποιείται χωρίς απαίτηση σύνδεσης (connectionless) μορφή σύνδεσης.

Ο τύπος της σύνδεσης μεταφοράς (transport connection) καθορίζεται από την RTSP URL.

Σε αντίθεση με το HTTP, το RTSP εξυπηρετητή να στέλνει αιτήσεις στον πελάτη.

Ωστόσο κάτι τέτοιο υποστηρίζεται μόνο από συνεχείς συνδέσεις μιας και ο εξυπηρετητής μέσων δεν έχει σε αντίθεση περίπτωση ασφαλή τρόπο να προσεγγίσει τον πελάτη. Εκτός αυτού είναι και ο μοναδικός τρόπος με τον οποίο οι αιτήσεις από τον εξυπηρετητή μέσων στον πελάτη μπορούν να διαβούν τους φράκτες (firewalls).

#### 7.1.8.1 Pipelining

Ο πελάτης που υποστηρίζει συνεχείς συνδέσεις ή μορφή χωρίς απαίτηση σύνδεσης (connectionless mode) μπορεί να διοχετεύει τις αιτήσεις του χρησιμοποιώντας την τεχνική διασωλήνωσης (pipeline) δηλ. να στείλει πολλαπλές αιτήσεις χωρίς να περιμένει την απόκριση της καθεμιάς. Σε τέτοια περίπτωση ο εξυπηρετητής πρέπει να στείλει τις αποκρίσεις των αιτήσεων κατά την ίδια σειρά με την οποία αυτές παρελήφθησαν.

### 7.1.8.2 Αξιοπιστία και βεβαιώσεις λήψης

Οι λήψεις των αιτήσεων επιβεβαιώνονται από τον παραλήπτη εκτός κι αν αυτές αποστέλλονται σε ένα multicast σύνολο. Αν δεν υπάρξει καμία επιβεβαίωση λήψης τότε ο αποστολέας μπορεί να ξαναστείλει το ίδιο μήνυμα μετά την έλευση κάποιου συγκεκριμένου χρονικού διαστήματος που ισούται με το χρόνο της μεταφοράς του μηνύματος στον προορισμό και της επιστροφής του στον αποστολέα (round-trip time, RTT.) Αυτός ο χρόνος εκτιμάται όπως και στο TCP με μια αρχική τιμή των 500 ms (Braden, 1989). Κάποια υλοποίηση μπορεί να αποθηκεύσει προσωρινά την τελευταία μέτρηση του RTT ως την αρχική τιμή για χρήση της σε μελλοντικές συνδέσεις.

Στην περίπτωση που χρησιμοποιείται κάποιο αξιόπιστο πρωτόκολλο μεταφοράς για να μεταφέρει το RTSP δεν πρέπει σε καμία περίπτωση να ξαναμεταδίδονται οι αιτήσεις. Αντι η RTSP εφαρμογή πρέπει να βασίζεται στην αξιοπιστία που παρέχει το υποκείμενο πρωτόκολλο μεταφοράς.

Αν και το υποκείμενο αξιόπιστο πρωτόκολλο μεταφοράς όπως το TCP αλλά και η RTSP εφαρμογή ξαναμεταδίδουν αιτήσεις, τότε είναι πιθανό κάθε απώλεια πακέτων να προκαλέσει διπλή επαναμετάδοση. Τυπικά ο αποδέκτης δεν είναι σε θέση να εκμεταλλευτεί την επαναμετάδοση του επιπέδου εφαρμογής, μιας και η στοίβα μεταφοράς δε θα παραδώσει αυτήν την επαναμετάδοση προτού φτάσει η πρώτη απόπειρα στον αποδέκτη. Αν η απώλεια του πακέτου οφείλεται σε σύγκρουση τότε οι πολλαπλές επαναμεταδόσεις, σε διαφορετικά επίπεδα, θα χειροτερέψουν την κατάσταση των συγκρούσεων.

Για την αποφυγή του προβλήματος αμφιβολίας και της επαναμετάδοσης, χρησιμοποιείται η Timestamp επικεφαλίδα. Κάθε αίτηση μεταφέρει έναν αριθμό σειράς στο πεδίο CSeq της επικεφαλίδας, που αυξάνεται κατά ένα για κάθε διακριτή αίτηση που μεταδίδεται.

Αν η αίτηση επαναμεταδοθεί λόγω έλλειψης επιβεβαίωσης λήψης της, τότε πρέπει να φέρει τον αρχικό (αυθεντικό) αριθμό σειράς (δηλ. Να μην έχει αυξηθεί αυτός ο αριθμός). Τα συστήματα που υλοποιούν το RTSP πρέπει να υποστηρίζουν την μεταφορά του RTSP πάνω από το TCP, όπως επίσης πρέπει να υποστηρίζουν το UDP. Ο προκαθορισμένος αριθμός θύρας που χρησιμοποιείται από τον RTSP εξυπηρετητή είναι 554 τόσο για το TCP όσο και στην περίπτωση του UDP.

Κάποιος αριθμός RTSP πακέτων που προορίζεται για τον ίδιο προορισμό είναι δυνατό να πακεταριστεί σε μια μόνο χαμηλότερου επιπέδου PDU ή να ενθυλακωθεί σε ένα TCP stream. Τα RTSP δεδομένα μπορούν να διαστρωματοποιηθούν με RTP και RTCP πακέτα. Σε αντίθεση με το HTTP το RTSP μήνυμα πρέπει να περιέχει μια επικεφαλίδα μεγέθους του περιεχομένου, όποτε υπάρχει μέσα στο μήνυμα κάποιο φορτίο. Διαφορετικά το RTSP πακέτο τελειώνει με μια κενή γραμμή αμέσως μετά την τελευταία επικεφαλίδα του μηνύματος.

Σε αυτό το σημείο ολοκληρώνεται η γενική περιγραφή του RTSP διότι περαιτέρω εμβάθυνση οδηγεί αναπόφευκτα ανάλυση τεχνικών όρων και λεπτομερειών.

## 8. Πρωτόκολλο SIP (Session Initiation Protocol)

---

Από την έγκρισή του στις αρχές του 1999 ως επίσημο πρότυπο, το πρωτόκολλο έναρξης συνόδου (SIP) έχει κερδίσει τεράστια αποδοχή στην αγορά σηματοδοτώντας υπηρεσίες επικοινωνιών στο διαδίκτυο.[1]

### 8.1 Ιστορία SIP

Το SIP έχει την προέλευσή της στα τέλη του 1996 ως συστατικό του συνόλου "Mbone" πρωτοκόλλων. Το Mbone, ή η πολλαπλής διανομή σπονδυλική στήλη, ήταν ένα πειραματικό πολλαπλής διανομής δίκτυο πάνω από το δημόσιο Διαδίκτυο. Χρησιμοποιήθηκε για τη διανομή του περιεχομένου πολυμέσων, συμπεριλαμβανομένων των συζητήσεων και των σεμιναρίων, τις ραδιοφωνικές μεταδόσεις των διαστημικών αποστολών, και των συνεδριάσεων IETF. Ένα από τα ουσιαστικά συστατικά του ήταν ένας μηχανισμός που επιτρέπει στους χρήστες να ακούσουν μέσα από μια τρέχουσα ή μελλοντική σύνοδο πολυμέσων στο διαδίκτυο. Βασικά - ένα πρωτόκολλο έναρξης συνόδου, Κατά συνέπεια SIP γεννήθηκε.

Σαν εργαλείο Mbone (και ως προϊόν του IETF), το SIP σχεδιάστηκε με ορισμένες προϋποθέσεις. Πρώτα ήταν εξελιξιμότητα: Δεδομένου ότι οι χρήσεις μπόρεσαν να κινηθούν οπουδήποτε στο διαδίκτυο, το πρωτόκολλο έρπετε να λειτουργήσει στην εκτενή ζώνη από την πρώτη ημέρα. Οι χρήστες θα μπορούσαν να προσκληθούν σε πολλές συνόδους, έτσι το πρωτόκολλο που απατήθηκε να είναι κλιμακούμενο και στις δύο κατευθύνσεις. Μια δεύτερη υπόθεση ήταν συστατική επαναχρησιμοποίηση: Παρά την εφεύρεση νέων εργαλείων πρωτοκόλλου, εκείνοι που αναπτύχθηκαν ήδη μέσα στο IETF θα χρησιμοποιούνταν. Αυτός περιέλαβε πράγματα όπως MIME, URLs, και sdp (που χρησιμοποιήθηκαν ήδη για άλλα πρωτόκολλα, όπως το SAP. Αυτό οδήγησε σε ένα πρωτόκολλο που ενσωματώθηκε καλά με άλλες εφαρμογές IP (όπως ο Ιστός και το ηλεκτρονικό ταχυδρομείο).

Η διαλειτουργικότητα ήταν ένας άλλος βασικός στόχος, αν και όχι συγκεκριμένος για το SIP. Η διαλειτουργικότητα είναι στην καρδιά της διαδικασίας και της λειτουργίας IETF, ως φόρουμ που παρακολουθείται από τους εφαρμοστές και τους λειτουργικούς εμπειρογνώμονες που χρίζουν πραγματικά και επεκτείνουν τις τεχνολογίες που σχεδιάζουν. Σε αυτούς τους πρακτικά-προσανατολισμένους τυποποιητές, η αρχή KISS (Keep It Simple Stupid) ήταν ο καλύτερος τρόπος να βοηθήσει να εξασφαλιστεί ακρίβεια και διαλειτουργικότητα.

Παρά τις ιστορικές δυνάμεις της, το SIP είχε σχετικά αργή πρόοδο καθ' όλη τη διάρκεια του 1996 και 1997, τότε περίπου άρχισε να απογειώνεται το ενδιαφέρον για την τηλεφωνία Διαδικτύου. Οι άνθρωποι άρχισαν αν θεωρούν το SIP ως τεχνολογία που θα λειτουργούσε επίσης για VoIR, μόνο σε συνόδους Mbone. Το αποτέλεσμα ήταν μια ενταγμένη προσπάθεια προς την ολοκλήρωση της προδιαγραφής στα τέλη του 1998, και ολοκλήρωσης μέχρι το τέλος του έτους. Έλαβε την επίσημη έγκριση ως RFC (ο επίσημος όρος για πρότυπα IETF) το Φεβρουάριο και την έκδοση ενός RFC αριθμού, 2543, το Μάρτιο.

Από εκεί, η βιομηχανική αποδοχή του SIP αυξήθηκε εκθετικά. Η εξελιξιμότητα, και – το σημαντικότερο – η ευελιξία της απευθύνθηκαν στους φορείς παροχής υπηρεσιών και τους προμηθευτές που είχαν τις ανάγκες που ένα κάθετα ενσωματωμένο πρωτόκολλο, όπως το H.323, δεν θα μπορούσε να καλύψει. Μεταξύ των φορέων παροχής υπηρεσιών MCI οδηγήθηκε στην ευαγγελική δαπάνη. Καθ' όλη τη διάρκεια του 1999 και μέσα στο 2000, είδε την υιοθέτηση από τους περισσότερους σημαντικούς προμηθευτές, και τις ανακοινώσεις των δικτύων από τους φορείς παροχής υπηρεσιών. Η τεράστια επιτυχία επιτεύχθηκε στη διαλειτουργικότητα μεταξύ των προμηθευτών. Άλλοι οργανισμοί προτύπων άρχισαν να εξετάζουν το SIP επίσης, συμπεριλαμβανομένης της ITU και του ETSI TIPHON, IMTC, κοινοπραξία Softswitch, και JAIN.

Κοιτάζοντας προς τα εμπρός, το 2000 είναι ένα έτος στο οποίο τα πραγματικά δίκτυα SIP επεκτείνονται, ωστόσο να αναγγελθούν πραγματικά προϊόντα, και οι εφαρμογές και οι υπηρεσίες αρχίζουν να εμφανίζονται.

## 8.2 Λειτουργία SIP

Όπως το όνομα υπονοεί, το πρωτόκολλο έναρξης συνόδου (SIP) είναι για την έναρξη των συνόδων διαλογικών επικοινωνιών μεταξύ των χρηστών. Το SIP χειρίζεται επίσης τη λήξη και τις τροποποιήσεις των συνόδων. Το SIP πραγματικά δεν καθορίζει τι μια "σύνοδο" είναι αυτό περιγράφεται από το περιεχόμενο που φέρεται στα μηνύματα SIP. Το μεγαλύτερο μέρος του SIP είναι το μέρος έναρξης, δεδομένου ότι αυτό είναι πραγματικά η δυσκολότερη πτυχή. Ένας χρήστης έχει ένα PC στην εργασία, ένα PC στο σπίτι, και ένα τηλέφωνο γραφείων IP στο εργαστήριο. Μια κλήση για εκείνο τον χρήστη πρέπει να χτυπήσει σε όλα τα τηλέφωνα αμέσως. Επιπλέον, ο χρήστης μπορεί να είναι κινητός. Μια ημέρα στην εργασία, και την επόμενη ημέρα επισκέπτεται ένα πανεπιστήμιο. Αυτές οι δυναμικές πληροφορίες θέσης πρέπει να ληφθούν υπόψη προκειμένου να βρεθεί ο χρήστης.

Μόλις βρεθεί ο χρήστης που καλείται, το SIP μπορεί να εκτελέσει τη δεύτερη κύρια λειτουργία του - παραδίδοντας μια περιγραφή της συνόδου στην οποία ο χρήστης προσκαλείται. Όπως αναφέρεται, το ίδιο το SIP δεν ξέρει για τις λεπτομέρειες της συνόδου. Αυτό που το SIP κάνει είναι να μεταβιβάζει τις πληροφορίες για το πρωτόκολλο που χρησιμοποιείται για να περιγράψει τη σύνοδο. Το SIP κάνει αυτό μέσω της χρήσης των για πολλές χρήσεις επεκτάσεων ταχυδρομείου Διαδικτύου (MIME), που χρησιμοποιείται ευρέως στις υπηρεσίες Ιστού και ηλεκτρονικού ταχυδρομείου για να περιγράψει το περιεχόμενο (HTML, ήχος, βίντεο, κλπ.) . Το πιο κοινό πρωτόκολλο που χρησιμοποιείται για να περιγράψει τις συνόδους είναι το πρωτόκολλο περιγραφής συνόδου (SDP), που περιγράφεται στο RFC2327 μπορεί επίσης να χρησιμοποιηθεί για να διαπραγματευτεί ένα κοινό σχήμα για την περιγραφή των συνόδων, έτσι ώστε άλλα πράγματα εκτός από το SDP να μπορούν να χρησιμοποιηθούν.

Μόλις βρεθεί ο χρήστης και η περιγραφή συνόδου παραδοθεί, το SIP χρησιμοποιείται για να μεταβιβάσει την απάντηση στην έναρξη συνόδου (δεχτείτε, απορρίψτε, κ.λπ.). Εάν γίνεται αποδεκτή, η σύνοδος είναι τώρα ενεργή. Το SIP μπορεί να χρησιμοποιηθεί για να τροποποιήσει τη συνόδου επίσης. Το κάνει πολύ εύκολα - ο δημιουργός, στέλνει το ίδιο μήνυμα με το αρχικά, αλλά με μια νέα περιγραφή συνόδου. Για αυτόν τον λόγο, η τροποποίηση των συνόδων (που περιλαμβάνει τα πράγματα όπως την προσθήκη και την αφαίρεση των ακουστικών ρευμάτων, προσθήκη του βίντεο, που αλλάζουν οι codecs, της αναμονής και του mute) υποστηρίζεται εύκολα με το SIP, εφ' όσον μπορεί να τις υποστηρίζει το πρωτόκολλο περιγραφής συνόδου.

Τέλος, το SIP μπορεί να χρησιμοποιηθεί για να ολοκληρώσει τη σύνοδο (δηλ., να κλείσει το τηλέφωνο).

### 8.3 Τρόπος λειτουργίας SIP

Το SIP είναι βασισμένη στο παράδειγμα αιτήματος - απάντησης. Για να αρχίσει μια σύνοδο, ο επισκέπτης (γνωστός ως πελάτη πρακτόρων χρηστών, ή UAC) στέλνει ένα αίτημα (αποκαλούμενο INVITE), που απευθύνεται στο πρόσωπο που ο επισκέπτης θέλει να μιλήσει. Στ SIP, οι διευθύνσεις είναι URLs. Το SIP καθορίζει ένα σχήμα URL που είναι παρόμοιο με το δημοφιλές mailto URL. Εάν η διεύθυνση ηλεκτρονικού ταχυδρομείου του χρήστη είναι [jdrosen@dynamic-soft.com](mailto:jdrosen@dynamic-soft.com), το SIP URL θα ήταν `Sip:jdrosen@dynamicsoft.com`. Αυτό το μήνυμα δεν στέλνεται άμεσα στο αποκαλούμενο συμβαλλόμενο μέρος, αλλά μάλλον σε μια οντότητα γνωστή ως κεντρικός υπολογιστής πληρεξούσιου (proxy server). Ο proxy server είναι αρμόδιος για τη δρομολόγηση και την παράδοση των μηνυμάτων στο αποκαλούμενο συμβαλλόμενο μέρος. Το αποκαλούμενο συμβαλλόμενο μέρος στέλνει έπειτα μια απάντηση, που δέχεται ή που απορρίπτει την πρόσκληση, που διαβιβάζεται πίσω μέσω του ίδιου συνόλου πληρεξούσιων, στην αντίστροφη διάταξη.



Ένα πληρεξούσιο μπορεί να λάβει έναν μόνο INVITE αίτημα, και στέλνει περισσότερα από ένα INVITE αιτήματα στις διαφορετικές διευθύνσεις. Αυτό το χαρακτηριστικό γνώρισμα, αποκαλούμενο “καρφώνοντας με τη διχάλα,” επιτρέπει μια προσπάθεια έναρξης συνόδου να προσπελάσει πολλαπλές θέσεις, με τις ελπίδες της εύρεσης του επιθυμητού χρήστη σε μια από αυτές. Μια στενή αναλογία είναι η υπηρεσία γραμμών τηλεφώνων σε ένα σπίτι, όπου όλα τα τηλέφωνα χτυπούν ταυτόχρονα.

Στο παράδειγμα μας, ο επισκέπτης ([jdrosen@dynamicsoft.com](mailto:jdrosen@dynamicsoft.com)) επιθυμεί να κάνει μια κλήση στον [joe@columbia.edu](mailto:joe@columbia.edu). Ο Jdrosen στέλνει το SIP INVITE μήνυμα στο πληρεξούσιο για [dynamicsoft.com](http://dynamicsoft.com) (βήμα 1). Αυτό το πληρεξούσιο διαβιβάζει έπειτα το αίτημα έξω στην Κολούμπια, όπου φθάνει στον κεντρικό υπολογιστή [Columbia.edu](http://Columbia.edu) (βήμα 2). Αυτός ο κεντρικός υπολογιστής δεν είναι πραγματικά ένα πληρεξούσιο, αλλά μια παρόμοια συσκευή αποκαλούμενη επαναπροσανατολισμένος κεντρικός υπολογιστής (redirect server). Αντί της αποστολής των κλήσεων, ο redirect server ζητά από τον αιτούντα να έρθει σε επαφή με τον επόμενο κεντρικό υπολογιστή. Ο κεντρικός υπολογιστής [Columbia.edu](http://Columbia.edu) κοιτάζει επάνω σε Joe στη βάση δεδομένων του, και καθορίζει ότι σήμερα, ο Joe είναι σε σαββατικό στο [foo.com](http://foo.com). Επομένως στέλνει μια ειδική απάντηση, αποκαλούμενη επαναπροσανατολισμένη, στο πληρεξούσιο [dynamicsoft.com](http://dynamicsoft.com), καθοδηγώντας τον για να δοκιμάσει αντ’ αυτού στο [joe@foo.com](mailto:joe@foo.com) (βήμα 3).

Το πληρεξούσιο [dynamicsoft](http://dynamicsoft.com) ενεργεί έπειτα σε αυτήν την απάντηση, το οποίο σημαίνει ότι προσπαθεί άμεσα να έρθει σε επαφή με [joe@foo.com](mailto:joe@foo.com). Έτσι, του στέλνει INVITE στον κεντρικό υπολογιστή [foo.com](http://foo.com) (βήμα 4). Αυτός ο κεντρικός υπολογιστής συμβουλευεται τη βάση δεδομένων του (το βήμα 5), και μαθαίνει (βήμα 6) ότι ο Joe είναι πραγματικά στις πωλήσεις. Έτσι, κατασκευάζει ένα νέο URL, [joe@sales.foo.com](mailto:joe@sales.foo.com), και στέλνει INVITE στο πληρεξούσιο [sales.foo.com](http://sales.foo.com) (βήμα 7).

Το πληρεξούσιο για το τμήμα πωλήσεων πρέπει έπειτα να διαβιβάζει INVITE στο PC όπου ο Joe κάθεται αυτήν την περίοδο. Πώς σε ξέρει ποιο PC είναι ο Joe; Το SIP καθορίζει ένα άλλο αίτημα, αποκαλούμενο ΚΑΤΑΛΟΓΟ, ο οποίος χρησιμοποιείται για να ενημερώσει ένα πληρεξούσιο για μια σύνδεση διευθύνσεων. Σε αυτήν την περίπτωση, όταν άνοιξε ο Joe τον πελάτη SIP του PC του, ο πελάτης θα καταχωρούμε τη δέσμευση [sip:joe@sales.engineering.com](mailto:sip:joe@sales.engineering.com) σε [sip:joe@mypc.sales.foo.com](mailto:sip:joe@mypc.sales.foo.com). Αυτό θα επέτρεπε στο πληρεξούσιο να ξέρει ότι Joe είναι πραγματικά στο mypc, ένας συγκεκριμένος οικοδεσπότης στο δίκτυο. Οι συνδέσεις που καταχωρούνται μέσω του SIP αναζωογονούνται περιοδικά, έτσι ώστε εάν κρυσταλλωθούν στο PC, η σύνδεση να αφαιρείται.

Το πληρεξούσιο sales.foo.com συμβουλεύεται αυτήν την βάση δεδομένων εγγραφής, και διαβιβάζει INVITE σε [joe@mypc.slaes.foo.com](mailto:joe@mypc.slaes.foo.com) (βήμα 8). Αυτή η INVITE φθάνει έπειτα στο PC του Joe και μπορεί έπειτα να αποκριθεί σε αυτό (έτσι το πρότυπο αίτημα-απάντησης). Το SIP δίνει πολλές απαντήσεις, και αυτές περιλαμβάνουν την αποδοχή, απόρριψη, επαναπροσανατολισμός, πολυάσχολος, και τα λοιπά. Η απάντηση διαβιβάζεται πίσω μέσω των πληρεξούσιων στον αρχικό επισκέπτη (βήματα 9,10,11,12). Μια αναγνώριση στέλνεται (ένας άλλος τύπος αιτήματος, αποκαλούμενος ACK) στο βήμα 13, και η σύνοδος καθιερώνεται. Το MEDIA μπορεί έπειτα να ρεύσει (βήμα 14).

Το SIP είναι επίσης διαμορφωμένο μετά από το HTTP από πολλές απόψεις. Το HTTP είναι επίσης αίτημα-απάντηση. Το SIP δανείζεται ένα μεγάλο μέρος της σύνταξης και της σημασιολογίας από το HTTP. Η κειμενική μορφοποίηση μηνυμάτων, η χρήση των επιγραφών, η υποστήριξη MIME, και πολλές επιγραφές είναι ίδιες. Ένας εμπειρογνώμονας HTTP που εξετάζει αν μήνυμα SIP θα είχε τη δυσκολία που τους διακρίνει.

## 8.4 Πλεονεκτήματα SIP

### Υπηρεσίες

Η τηλεφωνία Διαδικτύου άρχισε υπό την προϋπόθεση ότι ήταν φτηνότερη από την κανονική τηλεφωνική κλήση. Οι χρήστες ήταν πρόθυμοι να ανεχτούν την υποβιβασμένη ποιότητα ή τη μειωμένη λειτουργία για το χαμηλότερο κόστος. Εντούτοις, οι διάφορες δαπάνες εξαφανίζονται γρήγορα. Για να συνεχίσει να υπάρχει, η τηλεφωνία Διαδικτύου πρέπει να βρει έναν άλλο λόγο ύπαρξης. Η απάντηση είναι υπηρεσίες.

Μερικές από τις πιο συναρπαστικές εφαρμογές έχουν βρει ήδη τη θέση τους στο διαδίκτυο, αν και όχι (ακόμα) υπό μορφή υπηρεσιών πολυμέσων. Τώρα σκεφτείτε την ενσωμάτωση των επικοινωνιών πολυμέσων, όπως η φωνή, με τον Ιστό το ηλεκτρονικό ταχυδρομείο, τους καταλόγους φιλαράκων, το στιγμιαίο μήνυμα, και τα σε απευθείας σύνδεση παιχνίδια. Ολόκληρα τα νέα σύνολα χαρακτηριστικών γνωρισμάτων, υπηρεσιών, και εφαρμογών γίνονται κατανοητά.

Το SIP είναι ιδανικά ταιριασμένο εδώ. Η χρήση της URLs, υποστήριξης της για MIME και της μεταφορά του αυθαίρετου περιεχομένου (το SIP μπορεί να φέρει τις εικόνες, MP3s, ακόμη και applets της JAVA), και χρήσης των μηχανισμών δρομολόγησης ηλεκτρονικού ταχυδρομείου, σημαίνει ότι μπορεί να ενσωματωθεί καλά με αυτές τις άλλες εφαρμογές. Παραδείγματος χάριν, είναι εξίσου εύκολο να επαναπροσανατολίσει έναν χρήστη σε ένα άλλο τηλέφωνο όπως με το να επαναπροσανατολίσει έναν χρήστη ιστοσελίδας.

## **Εξελιξιμότητα:**

Το SIP χρησιμοποιεί το πρότυπο Διαδικτύου για την εξελιξιμότητα - γρήγορη και απλή στον πυρήνα, εξυπνότερο με το λιγότερο όγκο στην περιφέρεια. Για να ολοκληρώσει αυτό, το SIP καθορίζει διάφορους τύπους κεντρικών υπολογιστών πληρεξούσιου. Τα "Call-Stateful" πληρεξούσια ζουν γενικά στην άκρη του δικτύου. Αυτά τα πληρεξούσια ακολουθούν το κράτος κλήσης, και μπορούν να παρέχουν τα πλούσια σύνολα υπηρεσιών που εδρεύουν σε αυτήν την γνώση. Πιο κοντά στον πυρήνα, τα "transaction-stateful" πληρεξούσια ακολουθούν τα αιτήματα και τις απαντήσεις, αλλά δεν έχουν καμία γνώση της συνόδου ή του κράτους κλήσης. Μόλις γίνει αποδεκτή μια πρόσκληση συνόδου, το πληρεξούσιο την ξεχνά. «Όταν η λήξη συνόδου φθάνει, το πληρεξούσιο το διαβιβάζει χωρίς να πρέπει να ξέρει για τη σύνοδο.

Τέλος, τα "άνευ υπηκοότητας" πληρεξούσια υπάρχουν στον πυρήνα. Αυτά τα πληρεξούσια λαμβάνουν τα αιτήματα, όπως INVITE, και τα ξεχνούν αμέσως. Το πρωτόκολλο SIP παρέχει τις εγκαταστάσεις για να εξασφαλίσει ότι η απάντηση μπορεί να καθοδηγηθεί σωστά πίσω στον επισκέπτη. Τα άνευ υπηκοότητας πληρεξούσια είναι πολύ γρήγορα, αλλά μπορούν να παρέχουν λίγες υπηρεσίες. Τα Call-Stateful πληρεξούσια δεν είναι τόσο γρήγορα, αλλά ζουν στην περιφέρεια, όπου οι όγκο κλήσης είναι χαμηλότεροι.

## **Επεκτασιμότητα:**

Η ιστορία έχει διδάξει στους μηχανικούς Διαδικτύου ότι τα πρωτόκολλα επεκτείνονται και χρησιμοποιούνται με τρόπους που δεν σκόπευαν ποτέ (το ηλεκτρονικά ταχυδρομείο και ο Ιστός είναι κα τα δύο άριστα παραδείγματα αυτού). Έτσι, έμαθαν να σχεδιάζουν την υποστήριξη για το έκτατο εξαρχής.

Το SIP έχει τους πολυάριθμους μηχανισμούς για να υποστηρίξει τις επεκτάσεις. Δεν απαιτεί από τον καθέναν να εφαρμόσει τις επεκτάσεις. Οι εγκαταστάσεις που παρέχονται επιτρέπουν σε δύο συμβαλλόμενα μέρη να καθορίσουν το κοινό σύνολο ικανοτήτων, έτσι ώστε μια έναρξη συνόδου να μπορεί πάντα να ολοκληρωθεί.

## **Ευελιξία:**

Το SIP δεν είναι ένα πλήρες σύστημα για την τηλεφωνία Διαδικτύου. Δεν υπαγορεύει την αρχιτεκτονική τα σχέδια χρήσης, ή το σενάριο επέκτασης. Δεν εξουσιοδοτεί πόσοι κεντρικοί υπολογιστές υπάρχουν, πως συνδέονται, ή που κατοικούν.

Αυτό αφήνει στους χειριστές τεράστια ευελιξία στο πως το πρωτόκολλο χρησιμοποιείται και επεκτείνεται. Ένας τρόπος να σκεφτεί από το είναι ότι το SIP είναι ένας φραγμός LEGO, οι χειριστές μπορούν να συναρμολογήσουν μια πλήρη λύση με τη λήψη άλλων φραγμών LEGO, και την τοποθέτηση τους μαζί με τον τρόπο που βλέπουν ότι ταιριάζει.

## 8.5 Παροχή υπηρεσιών SIP

Πέρα από τη δυνατότητα του SIP να δημιουργηθούν οι νέες, καινοτόμες υπηρεσίες, το πρωτόκολλο έχει τη δυνατότητα για να καταστήσει τη δημιουργία υπηρεσιών διαθέσιμη στις "μάζες," με τον ίδιο σχεδόν τρόπο με το περιεχόμενο Ιστού. Αυτές οι υπηρεσίες μπορούν να δημιουργηθούν από τους φορείς παροχής υπηρεσιών, τους επιχειρηματικούς διοικητές, και τα τμήματα ΤΠ, ή ακόμα και άμεσα στο τέλος τους χρήστες. Με το άνοιγμα της καινοτομίας στο κοινό όλα τα είδη των νέων υπηρεσιών και τα χαρακτηριστικά γνωρίσματα μπορούν να αναπτυχθούν, δημιουργώντας ολόκληρες νέες αγορές. Έχουμε δει το ίδιο πράγμα στο διάστημα Ιστού - η δυνατότητα πρόσβασης του Ιστού έχει ενθαρρύνει μια τεράστια αγορά για το ηλεκτρονικό εμπόριο, δεδομένου ότι καθέναν με μια ιδέα και μερικά δολάρια μπορεί να βάλει μαζί το νέο περιεχόμενο. Στην υπερνίκηση του ιδιαίτερα συγκεντρωμένου και ελεγχόμενου από προτύπου της τηλεφωνίας, και στην τοποθέτηση των εργαλείων για τη δημιουργία υπηρεσιών σε τόσα πολλά χέρια, το SIP έχει εδώ τη δυνατότητα να παραδώσει το έξυπνο δίκτυο PSTN. Ποιο είδος υπηρεσιών ή εφαρμογών θα μπορούσε να επιτραπεί από το SIP; Εκτός από την παραδοσιακή κλήση-αποστολή, ακολουθώ, και ΜΗΝ-ΕΝΟΧΛΕΙΣ, το SIP έχει τη δυνατότητα για τη διευκόλυνση ολόκληρης νέας κατηγορίας υπηρεσιών που ενσωματώνουν τα πολυμέσα με τον Ιστό, το ηλεκτρονικό ταχυδρομείο, το στιγμιαίο μήνυμα, και "την παρουσία" (σημαμένος εδώ ως, "είστε αυτήν την περίοδο on-line;"). Η αξία που το Διαδίκτυο φέρνει στην τηλεφωνία Διαδικτύου είναι η ακολουθία των υπάρχουσων εφαρμογών που μπορούν να συγχωνευθούν με τη φωνή και τις τηλεοπτικές επικοινωνίες. Για παράδειγμα, στο τέλος μιας κλήσης, ένας χρήστης μπορεί να μεταφέρει το άλλο συμβαλλόμενο μέρος ιστοσελίδας αντί ενός άλλου τηλεφώνου. Αυτή η μεταφορά θα τελείωνε την κλήση, και θα ανάγκαζε τον ξεφυλλιστή Ιστού άλλου συμβαλλόμενου μέρους να πηδήσει στη νέα σελίδα. Στην ουσία, η αξία VoIP και SIP προέρχονται όχι από την ολοκλήρωση στο στρώμα δικτύων (δηλ., τρέξτε τις υπηρεσίες φωνής σας πάνω από το δίκτυο δεδομένων σας), αλλά στο στρώμα υπηρεσιών (δηλ., συνδυάστε τις υπηρεσίες φωνής σας με τις υπηρεσίες δεδομένων σας).

## 8.6 Ποιότητα Υπηρεσίας SIP

Ίσως το μεγαλύτερο πρόβλημα στο voice-IP, γενικά, είναι το ζήτημα της ποιότητας της υπηρεσίας. Η καθυστέρηση στις συνομιλίες που πολλοί χρήστες VoIP αντιμετωπίζουν προκαλείται από το jitter και τη λανθάνουσα κατάσταση της παράδοσης πακέτων μέσα το ίδιο στο Διαδίκτυο. Είναι χρήσιμο να θεωρηθούν μερικές από τις βασικές αρχές του Διαδικτύου για να καταλάβουμε τι μπορεί να γίνει για το πρόβλημα, ποια είναι η απάντηση του IETF και πώς προσκρούει στο SIP.[9]

Αυτήν την περίοδο, το Διαδίκτυο προσφέρει μια ενιαία υπηρεσία, παραδοσιακά καλούμενη "best effort". Με άλλα λόγια, όλα τα πακέτα είναι δημιουργημένα εξίσου. Δεν υπάρχει καμία διαφορά στο Διαδίκτυο εάν ένα πακέτο είναι ηλεκτρονικό ταχυδρομείο, FTP, ή download ιστοσελίδας. Εάν το Διαδίκτυο είναι πολυάσχολο, τα πακέτα χάνονται ή καθυστερούν.

Δυστυχώς, το ανθρώπινο αυτί είναι εξαιρετικά ευαίσθητο στη λανθάνουσα κατάσταση στην παράδοση του ήχου. Το ανθρώπινο αυτί μπορεί να ανιχνεύσει καθυστερήσεις 200 χιλιοστών του δευτερολέπτου ή μεγαλύτερες στις συνομιλίες φωνής.

Το SIP δεν αναμιγνύεται στην επιφύλαξη των πόρων δικτύων ή τον έλεγχο αποδοχής. Αυτό είναι επειδή τα μηνύματα SIP μπορούν να μην τρέχουν πέρα από τα ίδια δίκτυα που τα πακέτα φωνής διαπερνούν. Η πλήρης ανεξαρτησία της πορείας SIP και της πορείας φωνής επιτρέπει σε ASPs για να παρέχει τις υπηρεσίες φωνής χωρίς παροχή της συνδετικότητας δικτύων. Αυτό είναι ένα εξαιρετικά σημαντικό πλεονέκτημα της αρχιτεκτονικής SIP. Λαμβάνοντας υπόψη αυτό, το SIP στηρίζεται σε άλλα πρωτόκολλα και τεχνικές προκειμένου να παρασχεθεί η ποιότητα της υπηρεσίας.

Οι περισσότεροι χρήστες έχουν αναμετρηθεί με τα ζητήματα QoS είτε με την προσθήκη του εύρους ζώνης στα δίκτυα τους, είτε με την εφαρμογή των σύνθετων και ακριβών τεχνικών διαμόρφωσης, όπως το ATM, στην κυκλοφορία IP. Αυτό μπορεί να είναι λογικό για τις δια επιχειρηματικές VoIP διαμορφώσεις, δεδομένου ότι το δίκτυο μπορεί να αντιμετωπιστεί άμεσα.

Για να δημιουργήσετε QoS στο διαδίκτυο, πρέπει να δημιουργήσετε τις διαφορετικές κατηγορίες υπηρεσίας για τα πακέτα. Το IETF έχει υιοθετήσει δύο μεθόδους: Ο πρώτος είναι ενσωματωμένες υπηρεσίες (RFC2211 και RFC2212), επίσης γνωστά ως INTSERV. Ο δεύτερος είναι διαφοροποιημένες υπηρεσίες (RFC2475), ή DIFFSERV.

Αυτός ο χωρισμός της καθιέρωσης συνοδού και της επιφύλαξης QoS εισάγει μια ενδιαφέρουσα παρενέργεια: Κάποιος μπορεί να πετύχει (συγκεκριμένα, η οργάνωση κλήσης), ενώ άλλη (επιφύλαξη των πόρων) μπορεί να αποτύχει. Το αποτέλεσμα είναι ότι το τηλέφωνο μπορεί να χτυπήσει και να απαντηθεί ακόμα κι αν το δίκτυο δεν μπορεί να υποστηρίξει την κλήση. Για να χειριστεί αυτό το πρόβλημα, ένας μηχανισμός συζεύξεων έχει αναπτυχθεί για το SIP βασισμένη στην εργασία που γίνεται αρχικά μέσα στο PacketCable Forum Distributed Call Signaling(DCS) ε. Αυτή η σύζευξη επιτρέπει ότι στο SIP INVITE (συγκεκριμένα, sdp), να περιέχει τους δείκτες που λένε στον αποκαλούμενο χρήστη "να μην χτυπήσει το τηλέφωνο" έως ότου έχουν διατηρηθεί οι ικανοποιητικοί πόροι (χρησιμοποιώντας RSVP ή κάποιο άλλο μηχανισμό). Μόλις πετύχουν οι επιφυλάξεις, ο επισκέπτης στέλνει ένα νέο αίτημα, που μεταγλωττίζεται δοκιμαστικά "PRECONDITIONS MET", στον αποκαλούμενο χρήστη, δείχνοντας όροι οι πόροι είναι διαθέσιμοι, και το τηλέφωνο πρέπει να χτυπήσει. Φυσικά εάν η επιφύλαξη QoS αποτυγχάνει, η κλήση μπορεί προαιρετικά να συνεχιστεί με την καλύτερη προσπάθεια. Αυτό σημαίνει ότι τα συστήματα SIP μπορούν να συστήματα SIP μπορούν να χρησιμοποιήσουν τα περιεκτικά end-to-end πρότυπα QoS για την τηλεφωνία Διαδικτύου. Δεδομένου ότι το SIP το ίδιο δεν διευκρινίζει εκείνους τους μηχανισμούς, οι νέες και περιεκτικότερες υπηρεσίες QoS που ανακαλύπτονται μπορούν να χρησιμοποιηθούν χωρίς επιρροή του SIP.

## 8.7 Μελλοντικές κατευθύνσεις

Πού πηγαίνει από εδώ και πέρα το SIP; Το SIP θα κινήσει μέσω της διαδικασίας IETF, από προτεινόμενο τυποποιημένο RFC προς το επόμενο στάδιο, τα πρότυπα σχεδίων RFC. Σαν προτεινόμενο τυποποιημένο RFC, τα πρότυπα αρχίζουν να επεκτείνονται και να αποκτούν την εμπειρία εφαρμογής. Αμετάβλητα, τα προβλήματα και οι ασυνέπειες βρίσκονται, και αυτά διορθώνονται καθώς η έκδοση σχεδίων RFC χτίζεται. Η διαλειτουργικότητα κάθε χαρακτηριστικού γνωρίσματος πρέπει να καταδειχθεί και να τεκμηριωθεί, και δύο διαλειτουργικές εφαρμογές πρέπει να υπάρχουν.[9]

Οι πολυάριθμες επεκτάσεις είναι επίσης υπό ανάπτυξη. Αρκετοί από αυτούς πέφτουν κάτω από την ομπρέλα "του SIP-T," ή "του SIP για την τηλεφωνία," παρέχοντας έναν τρόπο για τα παραδοσιακά, σηματοδοτώντας μηνύματα του PSTN, όπως ISUP, για να φερθούν στα μηνύματα SIP. Αυτό επιτρέπει στο SIP να χρησιμοποιηθεί μεταξύ των πυλών ή softswitches με έναν τρόπο που παρέχει την πλήρη διαφάνεια διατηρώντας τις δυνάμεις και τις υπηρεσίες του.

Μια άλλη επέκταση στοχεύει στη σύνδεση του επισκέπτη με το σωστό τερματικό για έναν δεδομένο χρήστη. Οι αποκαλούμενες “προτιμήσεις επισκεπτών” και οι “ικανότητες caller,” ενσωματώνουν τις πληροφορίες παρουσίας (δηλ. κατάλογοι φιλαράκων) στα μηνύματα SIP. Ένας επισκέπτης Μπορεί να ζητήσει από μια κλήση για [joe@company.com](mailto:joe@company.com) να καθοδηγηθεί στο κινητό τηλέφωνο Joe, ή σε έναν πελάτη PC που υποστηρίζει το βίντεο.

Πολλή εργασία είναι επίσης υπό εξέλιξη στην υποστήριξη της υποδομής. Αυτό περιλαμβάνει μια βάση διοικητικών πληροφοριών SNMP (MIB) για τη διαχείριση των κεντρικών υπολογιστών, των πυλών, και των πελατών SIP, κώδικες επιλογής DHCP για το SIP να υποστηρίζει το autoconfiguration, και τους μηχανισμούς QoS που επιτρέπουν στις συνόδους SIP να πραγματοποιηθούν μόνο εάν το ικανοποιητικό εύρος ζώνης υπάρχει για να τους υποστηρίξει.

## 9. Βιβλιογραφία

---

1. Merilee Ford, H. Kim Lew, Steve Spanier, Tim Stevenson, Internetworking Technologies Handbook, Cisco Press, ISBN:1-56205-603-4
2. William Stallings, Fifth Edition, Data and Computer Communications, Prentice Hall International Editions, ISBN 0-13-571274-2
3. Daniel Minoli – Robert Keinath Artech House, Distributed Multimedia Through Broadband Communications Services ISBN 0-89006-689-2
4. Francois Fluckiger, Understanding Networked Multimedia, Prentice Hall, ISBN 0-13-190992-4
5. Guojun Lu Artech House Publishers, Communication and Computing for distributed multimedia systems, ISBN:0-89006-884-4
6. William Stallings, "High-Speed Networks, TCP/IP and ATM Design Principles", Prentice Hall, 1998
7. Shulzrinne, Casner, Federick, Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 1889, IETF, January 1996
8. Shulzrinne, Casner, "RTP Profile fro Audio and Video Conferences with Minimal Control", RFC 1890, IETF, January 1996
9. Shulzrinne, H., "Internet Services: From Electronic Mail to Real-Time Multimedia", KIVS'95 – Chemnitz, Germany, February 20-24, 1995